Pilot Implementation Study of

Risk-Limiting Audit Methods

in the State of Rhode Island

Report of the Rhode Island RLA Working Group

August 2019

Common Cause Rhode Island

verifiedvoting

BRENNAN CENTER FOR JUSTICE

Executive Summary

In October 2017, Rhode Island Governor Gina Raimondo signed into law a groundbreaking election security measure. Now, state law requires Rhode Island election officials to conduct risk-limiting audits, the “gold standard” of post-election audits, beginning with the 2020 primary. A risk-limiting audit (“RLA”) is an innovative, efficient tool to test the accuracy of election outcomes. Instead of auditing a predetermined number of ballots, officials conducting an RLA audit enough ballots to find strong statistical evidence that outcomes are correct. The law, enacted in the aftermath of two critical events relating to the 2016 elections, stems from decades of advocacy aimed at increasing the efficiency, transparency, and verifiability of political contests in the state. Rhode Island is now the second state, joining trailblazing Colorado, to mandate use of this modern tool statewide.

Following the law’s enactment, a group of professionals with expertise in election security and election administration formed the Rhode Island Risk-Limiting Audit (“RIRLA”) Working Group. As its name suggests, the RIRLA Working Group was established to assess the conditions in Rhode Island to help the state as it prepares to implement the law. The RIRLA Working Group recommended – and Rhode Island officials agreed – that the state should conduct pilot RLAs in advance of the 2020 deadline. The Rhode Island Board of Elections chose January 2019 as the date for the pilots and, based on several factors, selected Bristol, Cranston, and Portsmouth, Rhode Island as participating municipalities.

Leading up to the pilots, the RIRLA Working Group had regular conference calls, meetings, and other correspondence to gain greater familiarity with Rhode Island’s election laws, practices, and voting equipment. In partnership with the state, the RIRLA Working Group set a goal to plan and develop a trio of pilot audits that would both meet the state’s needs and adhere to the Principles and Best Practices for Post-Election Tabulation Audits. Ultimately, the RIRLA Working Group drafted three separate audit protocols, step-by-step instructions to guide those who would conduct the RLAs over the course of two days.

On January 16 and 17, the Rhode Island Board of Elections and members of the RIRLA Working Group met in Providence, RI to conduct the pilot risk-limiting audits. In conducting three unique RLAs – a ballot-level comparison, a ballot polling, and a batch comparison audits – the partnership sought to:

- **Familiarize** election officials with RLAs and provide them some comfort with conducting them through a hands-on learning experience;
- **Evaluate** Rhode Island’s election facilities, equipment, and other resources to determine their adequacy for administering RLAs;
- **Emulate** the actual environment and real-world circumstances in which the RLAs would be conducted, including by having officials manage unanticipated scenarios that could arise;
- **Time** the various audit steps and compile the data to compare the relative efficiency of the separate audit approaches;
- **Ascertain** any gaps or deficiencies that might hinder the initial rollout and implementation of RLAs or the state’s ability to conduct them beyond 2020;
- **Make a set of recommendations** to help officials conduct RLAs pursuant to state law and best audit practices.

This report tells that story. It details how, through a collaborative effort, Rhode Island successfully conducted three pilot RLAs. It provides a general overview of RLAs, including the advantages and disadvantages of each method. The report describes the history of election administration in Rhode Island, which has led to the environment in which elections are conducted in the state today. It also lays out the essential components of the audits – their design, software tools, and presentation – and provides the results of the audits. Finally, the report describes some of the key lessons learned throughout the process, and it makes recommendations, specifically that Rhode Island pursue ballot-level comparison audits, so that state officials and the public move towards RLAs seamlessly and to improve the state's experience with RLAs in the future.

This was a truly collaborative effort. It would not have been possible without countless hours of work from Miguel Nunez and Steve Taylor (Rhode Island Board of Elections); John Marion (Common Cause Rhode Island); Mark Lindeman and John McCarthy (Verified Voting); Wilfred Codrington III and Andrea Cordova (Brennan Center for Justice); Luther Weeks (Connecticut Voters Count); Ron Rivest, Mayuri Sridhar, and Zara Perumal (Massachusetts Institute of Technology); Suzanne Mello-Stark (Rhode Island College), independent volunteers Lynn Garland, Neal McBurnett, Tom Murphy, and many others who made the pilot audits and this report a success.

A host of recent events, including equipment malfunction, cyberthreats, maladministration, and human error, have undermined public confidence in American elections. Unfortunately, these types of occurrences are not likely to disappear. But the good news is that our officials can help, even in the face of constraints on their time and public resources. They can take steps both to decrease the number of incidents and to show that they are willing and able to address the problems as they arise. They can start by conducting RLAs, to assure the public that the *reported* winners of elections are the *actual* winners. Risk-limiting audits are an efficient, effective, and straightforward way to enhance public confidence in our elections that takes into account the realities of election administration. The success of the pilots in Rhode Island can and should serve as a model for what state and local officials across the country can accomplish, and how other individuals and organizations can provide valuable assistance. We hope to make that clear with this report.

Recommendations include:

- Implement a ballot-level comparison risk-limiting audit
- Establish objective criteria for which races will be audited
- Conduct a centralized audit
- Consult local election officials
- Conduct a practice audit
- Use Arlo audit software
- Appoint an ongoing expert advisory council
- Initiate rulemaking
- Develop schedule with milestones
- Endorse vendor recommendations

Part I:
Overview of Risk-Limiting Audits

Evidence-Based Elections and Risk-Limiting Audits: A Brief Introduction

The 2016 U.S. presidential election brought increased media and public interest in long-standing concerns about the accuracy of electronic vote tabulations, throughout the nation, and in Rhode Island specifically (see Risk-Limiting Audits in Rhode Island: The Background). Computer scientists have warned for many years that computerized voting and counting systems are vulnerable to error or malicious subversion, and must be checked using methods that do not rely on the correctness of hardware or software.[1] The U.S. intelligence community and other credible observers have reported on widespread cyberattacks on election systems during the campaign, including the data breach of a state voter registration database.[2] Officials emphasized that there was no evidence that any data had been changed, nor was there evidence that votes had been changed.

Unfortunately, due to poorly designed equipment and procedures, evidence that votes *hadn't* been changed was fragmentary. Tens of millions of Americans voted on systems that provide no verifiable record of their votes. Many others marked and cast their votes on paper ballots, but their states did not check any of these ballots against the official returns. Paper ballots and systematic comparisons of the ballots to official returns are prerequisites for evidence-based elections.[3] A voting system that may produce accurate results, but provides no way to know whether it did, is inadequate. It provides far too many ways for resourceful adversaries to undermine public confidence in election integrity.

The basic strategy for evidence-based elections can be summarized as follows: use paper ballots, protect them, and check them. More specifically:

1. Voters must vote on voter-marked paper ballots – either marked manually or using ballot marking devices, but in either case, with a convenient and accessible means for voters to verify their ballots and, when necessary, to mark replacement ballots before officially casting them. Direct Recording Electronic voting machines that produce "voter-verifiable paper audit trails" provide, at best, an obsolescent stopgap: most voters never check them, and often they are difficult to audit.

---

1. For an authoritative overview of these concerns see *Securing the Vote: Protecting American Democracy*, The National Academies of Sciences, Engineering, and Medicine, 2018, 5, https://www.nap.edu/read/25120/chapter/1.
2. For instance, see Adam Thorp, "Illinois election officials: 'Very likely' state was target of Russian hackers," *Chicago Sun-Times*, July 13, 2018, https://chicago.suntimes.com/news/illinois-election-officials-very-likely-state-was-target-of-russian-hackers/.
3. Philip B. Stark and David A. Wagner, "Evidence-Based Elections," *IEEE Security and Privacy*, 2012,10, https://www.stat.berkeley.edu/~stark/Preprints/evidenceVote12.pdf.

2. Voted paper ballots must be carefully stored and managed to ensure that no ballots are added, removed, or altered, and procedures should be established to provide strong evidence of proper ballot management. Security experts should review the means adopted to prevent ballot tampering, and compliance audits should be performed to confirm that those procedures are actually followed (e.g. that ballot containers were properly secured).

3. Voted ballots also must be checked in robust post-election vote tabulation audits, in which audit judges manually review a random sample of voted ballots (and possibly additional ballots) and compare them to the reported results – before the results are finalized. To the extent feasible, these audits should be risk-limiting audits (RLAs), which are very likely to correct any election outcome that is wrong due to mistabulation, by means of a full manual count. Both RLAs and recounts should rely on human inspection of the actual voted ballots, not on images or on copied ballots.

A risk-limiting audit provides a large, prespecified minimum chance, if a reported outcome[4] for an audited contest is incorrect (i.e., disagrees with what an accurate full manual count of the ballots would show), that it will be corrected through a full manual count. (Legally, the full manual count might be part of the audit, or it might be a separate recount required because of the audit findings.) The risk limit is the maximum chance that an incorrect contest outcome will not be corrected. For instance, an audit that has at least a 95% chance of correcting an incorrect outcome has a 5% risk limit. A RLA with a small risk limit that does not lead to a full manual count provides strong evidence that the reported outcome is correct (i.e. matches what an accurate full manual count would show).

Risk-limiting audits can be highly efficient: they can be designed to do the amount of work – no more and no less – required to confirm a particular election contest outcome (or multiple outcomes) to a prespecified risk limit. Risk levels for other contests can be measured simultaneously even if they do not achieve the pre-specified risk limit.

More and more states are requiring risk-limiting audits. In November 2017 Colorado conducted the first coordinated statewide RLAs, after a series of pilot audits dating back to 2010. In November 2018, Colorado conducted the first RLAs of statewide contests. Rhode Island has a statutory requirement to conduct RLAs beginning with the 2020 primary. State laws in California and Washington explicitly welcome pilot or voluntary RLAs, and local election officials in other states have conducted pilot audits on their own authority. And in 2017 the Virginia General Assembly amended its audit law, adding a requirement to conduct "risk-limiting" post-election audits annually, effective July 1, 2018.[5]

4. The outcome is the legal and official consequence of an election: for instance, who will take office, who will participate in a runoff election, or whether a referendum will pass.
5. Va. Code Ann. § 24.2-671.1 (2017). These audits do not fit the definition of RLAs used in this paper because they are not allowed to begin until after results are certified and the recount deadline expires.

RLA Methods

Overview

There are three major approaches to risk-limiting audits, each of which has variations and sometimes can be combined with the other approaches. The pilot explored all three of these approaches at a manageable scale in a centralized setting.

1. In **ballot-level comparison**, a random sample of voted ballots is manually interpreted, and each manual interpretation is checked against the machine interpretation of the same ballot.

2. In **ballot polling**, a random sample of voted ballots is manually interpreted, and the resulting manual vote counts are checked against the total machine counts to see if they provide strong statistical evidence that the reported outcome is correct. This method is very similar to exit polling.

3. In **batch comparison**, a random sample of "batches" is selected, and the votes in each batch are counted manually. A batch may consist of all the ballots cast in a precinct, or on a particular voting machine. These counts are compared to the corresponding machine or precinct counts, batch by batch, to measure discrepancies as in ballot polling or ballot-level comparison.

Each one of these approaches starts with generating random numbers and using these random numbers to determine the sample of batches or individual ballots to be examined. Sampled ballots are interpreted by hand and eye and compared with corresponding electronic results. All discrepancies are noted and test statistics are calculated to see whether the audit sample provides provide strong statistical evidence that the reported outcome is correct. If not, additional sample ballots are examined until the statistical evidence is sufficient or a full manual recount is ordered

State Administration of Risk-Limiting Audits

Until 2017, most pilot risk-limiting audits had been conducted for single local jurisdictions – usually counties. In 2017, Colorado became the first state to design and deploy a state-wide risk-limiting audit across all of its individual counties. Scaling up from counties to the state-wide level introduced a number of additional problems for both the theory and practice of risk-limiting audits, including the following challenges:

1. The necessity for development, deployment, documentation and instruction in using robust, enterprise-level software to manage sampling of ballots, entry of vote data for specified contests from sampled ballots, flagging of discrepancies, and calculation of risk levels as specified and coordinated at the state level and conducted in each individual county.

2. Mathematics and software to combine results from different risk-limiting audit methods used in different counties because of limitations of some vote tabulation systems (e.g., some counties had equipment that could imprint ID numbers of centrally counted ballots to facilitate ballot-level comparison methods, while others did not and thus had to use ballot polling methods).

3. Distributed training and supervision of permanent and temporary election officials across 64 counties spanning an area of over 104,000 square miles.

Rhode Island's small geographic area (1,214 square miles) and statewide uniformity of voting systems affords it a singular advantage of being able to choose whether to conduct risk-limiting audits either at one central location or dispersed across 39 individual municipalities or a few geographically distributed sites.

Any of the three major types of risk-limiting audits (ballot-level comparison, ballot polling or batch comparison) could be conducted at a central facility in Providence, so long as that facility has sufficient space to store approximately 1,200 ballot boxes on shelves (organized to facilitate retrieval of ballots) and space to conduct other audit operations (including re-scanning, retabulating and imprinting ballots, dividing ballots in boxes into smaller batches of ballots in folders, retrieval and counting of sampled ballots, and recording marks from selected contests for sampled ballots).

All three major types of risk-limiting audits also could be conducted primarily in geographically distributed locations, up to and including all 39 local municipal jurisdictions, so long as municipal election officials in each location have sufficient space and are willing to take responsibility for doing the audit work during the days following each election, and so long as the state can acquire, deploy, and use software that each municipality can run at their local offices, and that state election officials in Providence can use to manage and centrally coordinate audits across all 39 local municipalities – especially for multi-jurisdiction contests.

Pros and Cons of Each RLA Method

The three different risk-limiting audit methods tried in the pilot each have their own advantages and disadvantages which depend on such factors such as the type of voting equipment used, the number of contests being audited, and the margins of the contests.

1. Ballot-Level Comparison

Pros	Cons
This method typically permits RLAs to look at far fewer ballots than the alternatives – often under 100 ballots are needed to attain a 5% risk limit. Even when margins are close the number of ballots that need to be examined to achieve a risk limit is relatively easy to predict and smaller than in the other methods, unless the margin is so close that an entire recount is required. Since the manual interpretation of each ballot is directly compared to its machine interpretation, it also should be possible to identify the cause of any discrepancy. Finding the sources of discrepancies is key for continuously improving audits and election processes. Ballot-level comparison audits also support efficient, informative "opportunistic" auditing of all other contests that appear on the audited ballots but are not being audited to a pre-specified risk limit.	To conduct ballot-level comparison audits, there must be a way to match each ballot to its machine interpretation. Existing precinct-based scanners cannot establish this correspondence, because they do not store either the ballots or the machine interpretations in the order the ballots were cast. The ballots must be rescanned on equipment that keeps the ballots and interpretations (cast vote records) in order, preferably imprinting each ballot sheet with a unique ID that also appears in the corresponding cast vote record. Not all precinct-based scanners meet these two conditions. In that instance, each ballot must be rescanned on equipment that can establish this one-to-one correspondence. This form of audit is known as a transitive RLA.[6] To rescan all ballots expeditiously, some jurisdictions would need to purchase or lease additional systems that can imprint ballots and link the imprinted IDs to the new (rescan) machine interpretations. Also, rescanning all these ballots involves significant additional ballot handling with its associated burden of labor and chain-of-custody issues. The audit needs to be carefully conducted in a transparent manner so that observers can check the entire process, including the summation of the cast vote records to make sure the tallies match.

6. The name refers to the transitive property of equality: if A = B and B = C, then A = C. Here, if the original outcome equals the retabulation outcome, and if the retabulation outcome (according to strong audit evidence) equals what a full manual count would show, then the original outcome matches what a full manual count would show.

2. Ballot Polling

Pros	Cons
Ballot polling audits can be performed without rescanning any ballots regardless of the type of voting machines. Relatively few ballots need to be examined (although more than in a ballot-level comparison audit) as long as the margins are not close. Ballot polling can be used when ballots cannot be matched to their machine interpretations.	For close margins (less than 2%) the number of ballots that need to be examined expands dramatically. Even when the number of ballots to audit is not very large, boxes with voted ballots from most or almost all of the precincts may need to be opened. (See Table 1 on p. 54) The required workload is less predictable than in comparison audits, so an outlier sample can require lots of additional auditing even if the machine count was very accurate. This type of audit does not provide any information on the cause of any discrepancies. Ballot polling is not well suited to opportunistic auditing of local contests: without the ability to tell whether a particular ballot was tabulated correctly, small samples are not very informative.

3. Batch Comparison

Pros	Cons
Batch comparison audits are straightforward to administer and easy to understand. The methods used are similar to recounts conducted in some states, so election officials are often familiar with these methods. This method requires little special preparation in most jurisdictions, as ballots are generally stored by batches that can be compared to the initial subtotal for that batch and no additional scanning is needed. Also, a full manual count – which could be required if an RLA fails to find strong evidence for an election outcome – could be administered in much the same way as batch comparison, except that all the batches would be included. For smaller margins, the required counting steadily increases – but less rapidly than with ballot polling. Batch-level manual counts are relatively easy to conduct in parallel at multiple locations. This type of audit can provide information about the accuracy of particular machines and the accuracy of the process of auditing and reporting results. Investigating sources of discrepancies can be time consuming and difficult: the results of the manual audit must be checked to confirm any apparent discrepancy; and all the ballots in the batch must be evaluated to determine which ones may have caused the discrepancy. However, this method identifies many more differences than the other methods and can confirm consistent differences.	Batch comparison audits typically involve examining the largest number of individual ballots. This approach is not well suited to auditing local contests with a small number of batches: it would often require a full hand count or almost a full hand count.

Part 2:
Planning Rhode Island's Risk-Limiting Audit Pilots

Risk-Limiting Audits in Rhode Island: The Background

History of Rhode Island's Voting System

On January 16 and 17, 2019 the Rhode Island Board of Elections conducted three pilot risk-limiting audits with great success. It was the culmination of a two-decade journey toward more stringent verification of Rhode Island's elections.

To understand the context it is necessary to go back to the 1950s, when Rhode Island first used Shoup mechanical voting equipment. Those machines were durable, but provided no auditable paper trail. Because of their durability, the Shoup machines lasted until the 1990s. In 1994, then-State Representative James Langevin led a commission to look for Rhode Island's next voting system.[7] The commission pitted optical scan against direct-recording electronic (DRE) equipment, and ultimately recommended that the state adopt an optical scan voting system.

As the newly-elected Secretary of State, James Langevin purchased a new voting system in 1996, which featured paper ballots and Eagle OpTech ballot tabulators.[8] The purchase of new machines marked a return to the past because the use of paper ballots originated in Rhode Island in the 18th Century.[9] In 2016 the OpTechs were replaced with tabulators from Election Systems & Services (ES&S), as described in **Current Rhode Island Voting System**.

Rhode Island put numerous processes into state statute to accommodate the transition, including a requirement for logic and accuracy testing.[10] While the commission recommended a voting system that would allow "post-election assurance procedures such as retesting equipment and partial manual recount," neither the enabling statute nor rules adopted by the Rhode Island Board of Elections provided for implementing such procedures.[11] Meanwhile, other states were adopting paper ballots, and the majority of them required some type of post-election audit.[12]

7. *The New Ballot Box: A Plan to Bring Rhode Island's Election Process into the 21st Century*, The Commission to Study the Purchase of New Voting Equipment for Rhode Island, June 1994.
8. Tom Mooney, "Head of R.I. Elections Board says state is the last using 'Optech III P' voting machines and they don't meet federal standards," Politifact, March 16, 2013, https://www.politifact.com/rhode-island/statements/2013/mar/16/robert-kando/head-rhode-island-board-elections-says-state-last-/.
9. Suzanne Irene Mello, "A Detailed Forensic Analysis and Recommendations For Rhode Island's Present and Future Voting Systems" (PhD diss., University of Rhode Island, 2011), https://dfcsc.uri.edu/docs/Mello_Dissertation.pdf.
10. 17 R.I. Gen. Laws § 17-19-14 (2016)
11. *The New Ballot Box*, p. 6.
12. "Post-Election Audits," National Conference of State Legislature, last updated January 3, 2019, http://www.ncsl.org/research/elections-and-campaigns/post-election-audits635926066.aspx.

Risk-Limiting Audit Legislation

In the mid-2000's Rhode Island had a series of contentious recounts.[13] In their wake, the ACLU of Rhode Island introduced an omnibus election reform bill in 2009.[14] In 2013, Common Cause Rhode Island added language to the legislation, first drafted by Pam Smith of Verified Voting, to require the state to conduct pilot risk-limiting post-election audits.[15] For several years, the legislation stalled in the General Assembly,

The confluence of two events, both of which occurred in 2016, changed the fate of that legislation. First, there was foreign interference in the presidential contest between Hillary Clinton and Donald Trump. Although intelligence agencies did not find any evidence that Russia altered the vote tabulation, the events brought questions of election security to the forefront.[16]

Second, in Rhode Island a simple mistake caused election administrators to realize the value of post-election auditing. The default answers for ballot questions in the ES&S templates are "accept" or "reject." North Kingstown, Rhode Island had a long-standing preference for using "yes" or "no" as their preferred answers. At the time the policy was to maintain separate files for different ballot styles (mail ballots, precinct-cast ballots, sample ballots, ballots used for logic and accuracy testing). For the precinct-cast ballots the vendor changed the default answer per North Kingstown's request, but did not move the corresponding oval. However, the mark was moved on the ballots produced for logic and accuracy testing. The result was that on election day the DS200 tabulator was looking for the voter's mark in the wrong area.

On election night the unofficial count revealed an unusual result of the referendum: 5 "Yes" votes to 8471 "No" votes. This quickly raised suspicion among the public and election officials. In the days to follow, the Rhode Island Board of Elections ordered the ballot scanners to be reprogrammed and the ballot question recounted. In the end, the certified result – 9492 "Yes" votes to 4569 "No" votes – revealed the gross inaccuracy of the outcome that was initially reported.[17]

Using pilot audit language from the omnibus election administration reform bill, Common Cause Rhode Island drafted standalone RLA legislation in 2017. They secured the support of the Rhode Island Board of Elections, Secretary of State Nellie Gorbea, and the Rhode Island Town and City Clerks Association. Several compromises were made during the legislative process.

---

13. Talia Buford, "Court won't hear Alves' appeal," Providence Journal, October 17, 2008; Mike McKinney, "Lynch wins one primary," *Providence Journal*, October 23, 2008; Mike McKinney, "Alves' election request awaits," *Providence Journal*, October 9, 2008.
14. H 5326, 2009 Gen. Assemb., Jan. Sess. (Ri. 2009).
15. H 5660, 2013 Gen. Assemb., Jan Sess. (Ri. 2013); S 0421, 2013 Gen. Assemb., Jan Sess. (Ri. 2013).
16. "US report finds no direct foreign interference in 2018 vote," *Associated Press*, December 21, 2018, https://www.apnews.com/cd2618aaeb6040c5b57fb301361c76fd.
17. Donita Naylor, "Recount changes results of North Kingstown septic-loan ballot question from nay to yea," *Providence Journal*, November 16, 2016, https://www.providencejournal.com/news/20161116/recount-changes-results-of-north-kingstown-septic-loan-ballot-question-from-nay-to-yea.

First, it provided for phased implementation, with audits *allowed* to begin in 2018, but *required* beginning in 2020. This change was made to give election officials (who implemented a new voting system in 2016 and electronic pollbooks in 2018) adequate time to develop audit processes.

Second, audits are authorized for the primary elections, but not required. Rhode Island's September primary is among the latest in the United States. An audit that leads to a full hand count could potentially prevent the state from complying with the MOVE Act's requirement that military and overseas ballots be sent no later than 45 days prior to an election.

Finally, language was added to clarify that the audits would happen at the conclusion of any recounts.[18]

In May, 2017, John Marion sought and received advice about the draft legislation from the State Audit Working Group (SAWG) (see Appendix B). The SAWG also encouraged John Marion and the Rhode Island Board of Elections's Director and Deputy Director, Bob Rapoza and Miguel Nunez, to observe the nation's first statewide risk-limiting audit in Colorado in November 2017 and to meet with election officials and others in attendance.

MIT Professor and SAWG member Ron Rivest, and his students provided invaluable support by meeting with election officials to familiarize them with RLAs and different auditing methods.

On October 5, 2017, Governor Gina Raimondo signed H 5704 Sub A and S 413 Sub A into law, making Rhode Island the second state, after Colorado, to require risk-limiting audits. Unfortunately, the legislation did not come with an appropriation, or even a fiscal impact note, from the legislature. Fortuitously, in March 2018 Congress appropriated $380 million in funding for election security to the states via their existing Help America Vote Act (HAVA) accounts.[19]

Common Cause Rhode Island petitioned the Rhode Island State Board of Elections to set aside monies for the development and implementation of the risk-limiting audits.[20] On April 24, 2018 the Board voted to request that $400,000 of the $3 million that Congress appropriated to Rhode Island be allocated for the design and implementation of RLAs over the next five years.

18. The legislation also contained a scrivener's error, requiring audits for "statewide" instead of "state" elections, thus excluding state legislative contests from being audited. Legislation was introduced in 2018 and again in 2019 to correct the mistake.
19. Consolidated Appropriations Act, 2018, Pub. L. No. 115-141 (2018).
20. Letter sent to Rhode Island Board of Elections on March 26, 2018 (On file with author).

Current Structure of Election Administration in Rhode Island

Rhode Island's elections are administered by the Rhode Island Board of Elections, the Secretary of State and 39 local Boards of Canvassers.[21] Rhode Island has 421 precincts for general elections.[22] State statute caps the total number of registered voters at 3000 per precinct.[23] In the 2018 general election Rhode Island had 787,000 registered voters, and saw 381,272 ballots cast.[24] Of those ballots, 26,560 were centrally counted mail ballots. In 2011 Rhode Island amended the state's "emergency mail ballot" statute, making it easier for voters to request and vote a mail ballot in the 20-day period preceding an election. While not true early voting, that provision has led to a drastic increase in the use of mail ballots as de facto early voting.

The Rhode Island Board of Elections is a seven-member board appointed by the governor to staggered nine year terms. They direct a staff of twelve (12) full-time employees, as well as numerous seasonal employees during elections. They have plenary power over all election day administrative responsibilities including maintenance and deployment of voting equipment and ballots to polling places.

The Rhode Island Secretary of State is the state's Chief Election Officer under HAVA. The Secretary is elected to four-year terms. The Elections Division has a full-time staff of four employees. They are responsible for certifying candidates and questions and producing the ballots for Election Day.

Each of Rhode Island's 39 cities and towns has a three-person board of canvassers. Members are appointed on a partisan basis by the city or town council (two seats for the majority party, and one for the minority party). Staffing varies; in larger communities the boards have full-time employees, and in smaller communities the town clerk splits responsibility for election administration with other duties. They are responsible for, among other tasks, recruiting poll workers.

21. Rhode Island is an outlier because it does not have county-level government.
22. The Board of Elections has the ability to administratively reduce the number of precincts for primary elections.
23. 17 R.I. Gen. Laws § 17-11-1 (2012)
24. Ted Nesi, "RI voter turnout rose this year, but not to record level," *WPRI*, November 28, 2018, https://www.wpri.com/politics/ri-voter-turnout-rose-this-year-but-not-to-record-level/1624601098.

Current Rhode Island Voting System

By the early 2000s, Rhode Island's voting system was showing its age. The Optech Eagle scanners were breaking down at an alarming rate. After some controversy, the Rhode Island General Assembly passed a statute in 2015 that made it the Secretary of State's responsibility to purchase voting systems (until then, the Board of Elections had the duty).[25] In July 2016, Secretary of State Nellie Gorbea announced the purchase of the EVS 5.2.0.3 Voting System from Election Systems and Software (ES&S).[26] In addition to the ElectionWare software, the purchase included 590 DS200 digital ballot scanners for precincts and two DS850 high-speed scanners for central counting of mail and provisional ballots. Each precinct is allocated between one to three DS200 tabulators, depending on expected turnout. In 2018 Rhode Island purchased an additional 20 DS200s. (The state owns the DS200 tabulators and stores them in a central warehouse in Providence between elections.)

In addition to the other newly purchased voting equipment and software, Rhode Island continues to use ES&S's AutoMark ballot marking devices for accessibility. The state purchased the AutoMark devices prior to the 2006 elections using the some of its original HAVA funds.[27] The Board deploys one AutoMark per precinct.

The Rhode Island Board of Elections is responsible for all election systems, which come with an eight-year agreement for servicing by the manufacturer (which began on 7/01/2016). As part of the state's contract, ES&S provides an on-site contractor, Joe Vitale, who operates the DS850s.

Participating Cities and Towns in Rhode Island's Pilot RLAs

The Rhode Island Board of Elections choose three communities to participate in the RLA pilots – Bristol, Cranston and Portsmouth, Rhode Island – based on several different factors, including variation in size and election administration staffing.

Bristol, Rhode Island is a town located on the eastern side of Narragansett Bay. The town, which is best known as the home of America's oldest Fourth of July parade, has 22,290 residents and 16,357 registered voters. The Town Clerk, currently Louis Cirillo, is elected.

Cranston, Rhode Island is the state's second largest city. Along with neighboring Warwick, Rhode Island, Cranston hosts the Gaspee Days, which commemorates the 1772 burning of the HMS Gaspee, a British revenue schooner. Cranston has 81,202 residents and 57,380 registered voters. The City Registrar, currently Nicholas Lima, was appointed by the

25. Jennifer Bogdan and Katherine Gregg, "Political Scene: R.I. voting-machine upgrade under new jurisdiction," *Providence Journal*, July 12, 2015, https://www.providencejournal.com/article/20150712/NEWS/150719783/13943.
26. Office of the Rhode Island Secretary of State, "Rhode Islanders welcome state-of-the-art voting systems," press release no. 28126, Jul 21, 2016, https://www.ri.gov/press/view/28126.
27. *Rhode Island State Plan: Help America Vote Act of 2002 (HAVA), Second Revised and Updated Version, Office of the Secretary of State*, 2010, http://www.elections.ri.gov/publications/Election_Publications/Voter_Info/FinalStatePlan2010.pdf.

Portsmouth, Rhode Island is at the northern end of Aquidneck Island. Founded in 1638 by Ann Hutchinson after her banishment from the Massachusetts Bay Colony, the town has 17,389 residents and 14,539 registered voters. The Registrar of Voters, currently Jacqueline Schulz, was appointed by the Board of Canvassers.

Rhode Island RLA Working Group

A group of activists and advocates reached out to the Rhode Island Board of Elections in 2018 with the offer to assist in the design and implementation of RLAs. That group calls itself the Rhode Island RLA Working Group (RIRLA Working Group) and is modeled after a similar effort in Colorado.

Quickly the RIRLA Working Group decided that a series of pilot RLAs were the appropriate first step in implementing the new law. The RIRLA Working Group recommended that the pilots prioritize collecting detailed data on the timing and staff required to carry out three major types of risk-limiting audits to develop RLA cost estimates for the state.

The Rhode Island Board of Elections, including its Director Bob Rapoza, agreed to the pilot, and Deputy Director Miguel Nunez and his staff (particularly Steve Taylor) became active participants. The Rhode Island Board of Elections set January 15-17 as the dates to conduct the pilot RLAs.

To help implement the pilot audits in January 2019, a number of volunteers as well as paid staff from participating organizations served on several sub-groups to plan and carry out the work, including:

- Overall planning, scheduling, and logistics – **Miguel Nunez** (RI Deputy Elections Director), **Steve Taylor** (RI Elections Department), **John Marion** (Common Cause Rhode Island)
- RLA Pilot Audit Design – **Mark Lindeman** (Verified Voting)
- Municipality selection – **Miguel Nunez** (RI Deputy Elections Director), **John Marion** (Common Cause Rhode Island), **Mark Lindeman** (Verified Voting)
- Software development, support, and coordination - **Mark Lindeman** (Verified Voting), **Tom Murphy**, **Ron Rivest** (MIT and Verified Voting), **Mayuri Sridhar** (MIT student), **Zara Perumal** (former MIT student), **Neal McBurnett**, **Suzanne Mello-Stark** (Rhode Island College)
- Timing Measurements – **Lynn Garland**, **Luther Weeks** (Connecticut Voters Count), **John McCarthy** (Verified Voting volunteer), **Mark Lindeman** (Verified Voting)
- Batch Counting RLA Method – **Luther Weeks** (Connecticut Voters Count)
- ES&S Vendor Discussions – **Wilfred Codrington** (Brennan Center), **Lynn Garland**
- RI RLA Group Moderation & Coordination – **John McCarthy** (Verified Voting volunteer)

Other regular contributors to RIRLA Working Group weekly teleconferences and lively email discussions included:

- **Andrea Cordova** (Brennan Center)
- **Ray Lutz** (Citizens Oversight)

We also had key contributions and intermittent participation from:

- **Dwight Shellman** (Colorado Department of State County Coordinator, who has spearheaded statewide risk-limiting audit efforts there since 2016)
- **Jerome Levato** (Election Assistance Commission, who formerly worked with Dwight Shellman to develop Colorado's statewide risk-limiting audits)
- **Brenda Cabrera** (City of Fairfax, Virginia General Registrar, who carried out a pilot RLA in August 2018)
- **Jennifer Morrell** (Democracy Fund and former Elections Director who piloted and implemented risk-limiting audits in Arapahoe County, Colorado)
- **Liz Howard** (Brennan Center and former Deputy Commissioner for the Virginia Department of Elections)
- **Tom Ryan** (Pima County, Arizona computer scientist and election integrity advocate)

Pilot Design

The pilot audits were designed to learn as much as possible about conducting RLAs in Rhode Island, while considering specific constraints created by resources and the state's voting system. Accordingly, the RIRLA Working Group made the following pilot audit choices:

The pilot would experiment with three separate RLA methods. To provide the full range of options for Rhode Island the RIRLA Working Group decided to pilot all three methods of RLAs. Because of Rhode Island's small size and large margins in most statewide contests, it was possible to pilot all three.

The audits would be limited to federal and statewide contests. Statewide and federal contests do not have some of the idiosyncrasies of local contests (e.g. "Vote for Five") that would complicate the audit design. Furthermore, because some local clerks are elected, limiting the audited contests to federal and statewide contests avoided having someone involved in the planning and implementation of an audit of their own election.

The samples would be taken from just three jurisdictions. Because ballots are stored by municipalities, it was crucial to have their cooperation in planning and implementing the audits. The pilots sought to confirm the correct outcome for the statewide and federal contests within the jurisdiction, not statewide. That reinforced that the point of the pilots was not to correct an incorrect outcome, as it would be in an actual RLA. (In a full risk-limiting audit, which would be conducted before the election results were certified, statewide contests would be audited using statewide samples, and it would not matter who "won" a particular municipality.)

The sample size would be predetermined. A full risk-limiting audit calls for the ballot sample size to be expanded until the risk limit is attained. For the pilots, however, we decided in advance that ballot sample sizes would not be expanded regardless of the results of the audit. This allowed for the RIRLA Working Group to better predict the time needed to complete the pilots.

We chose Bristol, Rhode Island for the ballot-level comparison method. This method required all voted ballots for the race or question being audited to be retabulated and imprinted. Bristol voters cast approximately 9,000 ballots, a number that was small enough to make retabulation and imprinting manageable, but large enough to teach participants the logistics of retabulation and imprinting. (Bristol's small number of cast mail ballots were excluded for the purposes of simplifying the audit design.)[28] We wanted the sampling design to: 1) meet stringent risk limits for most or all contests audited; 2) allow variation in the number of contests audited per ballot (so as to make time-based comparisons); and 3) be completed in approximately two hours. *Notably, for the ballot-level comparison method, we:*

28. There were 67 provisional and manual-count ballots in Bristol, about 0.7% of all ballots. These ballots could not be incorporated in the retabulation without some danger of compromising voter privacy. An official risk-limiting audit would have to take account of all voted ballots, either by auditing them or by making worst-case assumptions about any ballots excluded from the audit. (Here, even worst-case assumptions would be unlikely to affect the audit at all, because the fraction of ballots excluded was so small.) For simplicity – bearing in mind, again, that we were verifying municipal "outcomes" with no statewide consequence – we decided to treat the retabulation results as complete, thereby excluding these ballots from the pilot entirely.

- *Decided on a nested two-part sample, using a total of 100 ballots*, where Question 2 (see Appendix P-R) would be audited on all 100 ballots and the remaining federal and statewide contests would be audited on the last 50 ballots. We made this choice because Bristol's 11.9% diluted margin for Question 2 and 18.4% diluted margin for the governor's race meant that a sample of:
  - 64 ballots or fewer was likely to attain a 5% risk limit for state Question 2;
  - 41 ballots or fewer was likely to attain a 5% risk limit for the governor's contest; and
  - A smaller sample was likely to attain a 5% risk limit in the eight other contests.
- *Retabulated the ballots the week prior to the public portion of the pilot.* While we had considered completing the retabulation during the public portion of the pilot, conducting it before the audit gave us an opportunity to adapt to any unexpected events stemming from the retabulation results (e.g. unanticipated idiosyncrasies in the format of the data output). Instead, we presented a video that documented the retabulation process.

We chose Portsmouth, Rhode Island for the ballot polling method. Portsmouth voters cast almost 8,000 ballots. This was a large enough number to ensure that the audit could be conducted using a relatively small proportion of all ballots. Yet it was small enough to limit the amount of work both to divide the ballots into batches and determine the number of ballots per batch (see Part III). The audit could have treated each precinct's in-person ballots as a batch, and it could have relied on the Election Certificates submitted by pollworkers on election night for the number of ballots contained in each batch. (All mail ballots from Portsmouth would be treated as one batch.) However, we were reluctant to work with batches that could contain in excess of 1,500 ballots apiece and might be stored in multiple containers. For the pilot, we divided most precincts into smaller batches of varying sizes, to see how batch size affected the time required to retrieve ballots included in the sample. We also obtained independent counts of the total ballots in each batch rather than rely upon the precinct-based scanners and the electronic pollbooks (as the Election Certificates do). For simplicity's sake, we decided to audit only one contest using the ballot-polling method. We also fixed the ballot sample size at 200 ballots. Ultimately, we chose to audit the governor's race because the contest was fairly high on the ballot and was somewhat competitive (but not so competitive that 200 ballots would be unreasonably small in the jurisdiction). The simulation study showed that if Portsmouth's reported 22.2% margin for Gina Raimondo was roughly accurate, the audit sample had approximately an 84% chance of attaining a 10% risk limit.

We chose Cranston, Rhode Island for the batch comparison method. Specifically, the audit tested the in-person voted ballots from two precincts: 0704 and 0718. Precinct 0704 had about 230 in-person ballots, and 0718 had just over 600 in-person ballots.[29] Selecting two precincts with a relatively small number of ballots allowed participants and observers to compare two counting methods, the sort-and-stack and the hashmark. The contests for Senate, House of Representatives, and Governor were audited using the batch comparison method.

29. Mail ballots were excluded from this part of the pilot because retrieving mail ballots from just these two precincts would have been laborious. If Rhode Island chose to use batch comparison for statewide audits, it still could use ballot-level comparison for the mail ballots (as mentioned earlier), or it could treat the mail ballots from each municipality as a batch, which is how they are stored. The latter approach is most likely for a full manual count.

Audit Software

The Rhode Island risk-limiting audit pilots used the auditing software tool named Audit Conductor.[30] The software performed three principal functions: it identified and generated the pull lists for the randomly-selected ballots (or batches of ballots), reported the current state of the audits, and computed the risk levels.

The software tool is used to speed up the process of conducting the RLA, but is not a necessary component. Risk-limiting audits can be replicated by anyone, even without software, as long as they have the right information, including:[31]

- The random seed (see Part III for explanation);
- The cast vote record (CVR) file used in the audit, if any;
- The pre-audit reported election outcomes; and
- A record of the ballot interpretations from the audit.[32]

Development of the software tool started under the supervision of Suzanne Mello-Stark as a project to implement common risk-limiting methods (ballot polling and ballot-level comparison) in code. It was generalized, under the leadership of Tom Murphy, to be a simple, flexible, easy-to-use tool that could integrate existing risk-limiting tools.[33] Development was done publicly on GitHub, and the collaborators' main interactions were online and via conference calls. Additional code to process and format data, not part of Audit Conductor, was written by Mark Lindeman.

In the development of Audit Conductor three primary factors were considered; flexibility, simplicity and ease-of-use:

Flexibility

Audit Conductor was built with flexibility as a top concern. Given that a goal of the RLA was to compare several methods for conducting audits, software flexibility was especially important. In particular, we needed to have the capability to add methods as needed. To do this, we used Python as an implementation language because it has become a "lingua franca" for RLA audit code. By choosing the language that developers frequently use to write audit code, we were able to integrate tools from several "best-in-class" codebases into our own.

30. For audit software tool, see "Audit-Conductor," Election Audit Ware, GitHub, https://github.com/ElectionAuditWare/audit-conductor.
31. See details at Public RLA Oversight Protocol: http://bcn.boulder.co.us/~neal/elections/PublicRLAOversightProtocol.pdf.
32. Note that, for purposes of transparency and software-independence, ballot interpretations were recorded both on paper and in the audit software.
33. These tools included RIWAVE and BCTool (implementations of auditing statistical methods), and the rivest-sampler package to get a reproducible random sample.

Simplicity

Audit Conductor was also designed to be as simple as possible while fulfilling its requirements; namely generating pull lists, reporting on the current state of the audits, and computing the risk levels. The software's simplicity had two main benefits: it made it readable and it allowed us to delegate the more complicated statistical work to experts.

Because the Audit Conductor's source code is readable, i.e. others can understand it easily, the program can be audited for bugs. Having the code reviewed by outside observers (which, in the case of an open-source project like Audit Conductor, can be virtually any software developer) gives more confidence.

To build a tool that uses various sampling and audit methods, it is important that a team with limited resources be able to delegate specialized work to experts. Instead of reinventing the wheel – and spending time verifying that "the wheel" had been precisely reinvented – we used, in many cases, an expert-written, "off-the-shelf" tool. We could then focus our efforts on developing and testing the code unique to Audit Conductor.

Ease-of-use

The user interface was designed to be not only simple, but also intuitive and unambiguous for those conducting the audit. We strived to incorporate visual "checks" for each action taken by a user, to catch mistakes while the action was in-progress. For example, after clicking a radio button, the user would observe that the full selection is highlighted in a different color, which decreases the risk of inadvertently making the wrong selection. In addition, after entering selections on a particular ballot, users are presented with a confirmation screen to verify their selections.

Future Work

Audit Conductor satisfied the needs of the Rhode Island RLA pilot, but given the limited resources and the short timeline, it is not currently robust enough to use in an actual RLA. For example, it needs the ability to import the native ES&S CVR format or simultaneous data entry by multiple audit boards. In addition, for future use Audit Conductor can use improvements to its reporting functions, test suite, documentation, and ease of installation. In particular, cross-boundary tests written by statistical experts would be quite useful. A small bug surfaced during the audit,[34] and though it was quickly corrected, it would have been caught by these types of integration tests.

34. Specifically, one software module expected a CSV file to have its first row be a header, while another expected there to be no header. This caused mismatches in ballot comparison, where indexes were "off by one." Because all entered data was logged, the software team was able to easily re-run the code after fixing the error (by simply adding 1 to the index to account for the header row) and obtain the correct results.

Risk-Limiting Audit Logistical Preparation

The Rhode Island Board of Elections (RIBOE) is required by statute to implement risk-limiting audits beginning with the April 2020 primary. For the January 2019 pilots it was important to determine what physical assets and human capital the RIBOE must deploy. That includes where the audits are conducted and who does the work of conducting the audits.

Location Decision

The RIBOE has a 10,000 sq. ft. warehouse located in the basement of its main office in Providence, Rhode Island. The warehouse is currently used to store some of the equipment, including the DS200s and DS850s. It has several secure spaces that are used for processing mail ballots during elections. Ballots are stored locally by the state's 39 cities and towns, but all are within a 60-minute drive of Providence, except New Shoreham (Block Island).

The RIBOE warehouse was chosen as the site for the pilots because of its central location, size, security, and flexible floor plan.

Floor Plan

Part of the pilot audit planning included creating a floor plan (see Appendix D) for the various processes involved. The audits took place in two sections of the warehouse: on the main floor, where approximately 2500 square feet were utilized, and in a secure mail ballot processing room, where approximately 1150 square feet were used.

The main floor was used largely for two functions. First, this is where the more than 60 observers present for the pilot audits were seated during the introductory remarks, the dice roll to create the random seed, the manual examination of votes in the ballot-level comparison and ballot-polling audits (when they wished to watch), and the concluding panel. To facilitate observing the manual examination of ballots, the RIBOE set up two large screens and an audio amplification system. One screen displayed a projection of the voted ballot that was interpreted by the RIBOE staff and was confirmed or corrected by the pair of election judges. The other screen showed the user interface for the audit software tool as other RIBOE staff entered the vote.

The mail ballot processing room was used for the entire batch comparison audit as well as ballot retrieval for the ballot-level comparison and ballot polling audits. It also served as the secure storage location for the ballots throughout the two days. The processing room was set up with 8 tables each with 5 chairs for the teams of temporary employees recruited, trained, and paid by the RIBOE to conduct portions of the three pilots occurring in that area.

Expenses

The RIBOE had three sources of human capital for the pilot RLAs. First were the members of the Audit Work Group as described above. A second was the full-time staff of the RIBOE including Miguel Nunez, Steve Taylor, Jennifer Regan, and Manuel Hernandez. Each performed a variety of tasks during the preparation and execution of the audits. The final group were 18 temporary employees recruiting from a pool of workers normally employed by the RIBOE for mail ballot processing. Those employees were paid $100 per day for two days.

In addition to the human capital, the RIBOE had to purchase or rent the following equipment and services:

Item	Cost
Red & Blue ink cartridges for the DS850 *(we used the blue cartridge)*	$14.95/per
20 colored 10-sided dice	$25.90
PA system *(we upgraded our portable sound system)*	$1474.48
Document Camera	$199.00
One counting scale (*Mark Lindeman also loaned two additional scales)*	$295.00
6 combination squares (instead of rulers)	$7.96/per
Several different types of ballot storage boxes	$24.00
Videographer	$3900.00
Rental of Chairs and table skirts *(for lunch and dinner for those present)*	$938.00
Food and drink	$1341.12

In total, the two-day pilot RLAs cost the RIBOE $12,705.06. Those costs do not include staff and volunteer time, or assets already owned by the RIBOE such as projectors, screens, etc. Going forward the bulk of the cost for RLAs in Rhode Island will be staff time and the cost of any temporary employees needed.

Part III:
Rhode Island's RLA Pilot Experience & Results

Random Sampling and Ballot Manifests

RLAs crucially depend on valid random samples of ballot sheets[35] (or batches). In most cases, including the Rhode Island pilot, these samples are Simple Random Samples: every ballot is equally likely to be selected. The random sampling method used should be observably fair – not subject to manipulation or even inadvertently biased toward including some ballots rather than others – and not predictable before the sample is selected.

In the typical RLA random sampling method, observers roll a ten-sided die twenty times, creating a string of numbers known as a random seed. That string is typically entered into a software tool. The software tool uses a pseudorandom number generator, with the seed, to generate a sequence of random numbers. Then the tool maps those random numbers to a list of the stored ballots known as a ballot manifest in order to select the ballots (or batches) to be audited. To draw a sampled ballot, one needs to know how many ballots there are, and exactly where they are stored. A ballot manifest serves this purpose: it accounts for every ballot and says where to find each one. When carefully designed, this approach provides assurance that the sample could not have been predetermined or known in advance, before the seed was created – but also allows people to verify that the sample is correctly derived from the seed. Ideally, the ballot manifest should be prepared without relying on the voting system itself to correctly report how many ballot sheets are in each batch.

35. Long ballots often are divided among two or more pieces of paper, or sheets. All jurisdictions in the pilot used single-sheet ballots, so most parts of this report use "ballots" and "ballot sheets" interchangeably. However, this distinction becomes important when planning for future audits that include multiple-sheet ballots.

Audit Design Specifics

Each type of RLA method piloted required a specific design tailored to Rhode Island-specific factors. For detailed protocols see Appendix E. This section explains the key elements in the design for each method.

1. Ballot-Level Comparison

The basic steps in the ballot-level comparison audit were as follows:

- Retabulate the ballots, imprinting them with unique sequential ID numbers. (This step was done the week before the public pilot began.)
- Divide the Bristol ballots into batches and prepare a ballot manifest accordingly.
- In a public ceremony, use dice to generate a random seed; use this seed entered into the Audit Conductor software to select a random sample of 100 ballots.
- Retrieve the ballots in the sample, using the imprinted IDs.
- Interpret selected contests on the ballots in the sample and enter into Audit Conductor.
- Examine the results produced by the Audit Conductor software and investigate any discrepancies.

Retabulation and Batching

The retabulation of Bristol ballots was conducted on January 8, eight days before the public pilot began. Mark Lindeman observed the retabulation and participated in discussions as it was conducted; John Marion also observed most of the retabulation. A timekeeper captured timing data for each precinct; the timing results are discussed later in the report. The basic workflow was as follows, proceeding serially as in an assembly line:[36]

1. Open ballot containers and manually orient ballots face up and top up.
2. Divide ballots into batches, using a counting scale to estimate the number of ballots per batch as one source of data for the ballot manifest.
3. Retabulate each batch of ballots on the DS850 scanner, imprinting ballots with a serial ID number.
4. Reconcile the scanner count of the number of ballots with the scale count; place the batch in a folder, assigning a batch ID; complete the ballot manifest for the batch.
5. Store and seal the ballots in a new container with the batch ID.

Some details deserve further discussion.

Step 1. Orienting the ballots “right-side-up” was not strictly necessary, but it seemed likely to help staffers retrieve ballots by imprinted ID during the public pilot, because it ensured that most IDs would be in the same corner of the ballot. Staff were available for this task, and it was not a bottleneck. (Generally three election workers performed this step.)

36. During the retabulation, one batch was scanned out of order, before the previous precinct was complete. We decided to delete the batch’s retabulation data and rescan it after completing the previous precinct, so the imprinted ID numbers used in the audit would be in consistent order.

Step 2. Two counting scales were tested in the pilot: a Scalemart CS-20 scale with 6-pound capacity and 0.0002-pound reporting precision, and a Tree (LW Measurements) LCT-16 scale with 16-pound capacity and 0.0005-pound precision. A stack of 200 (later 300) ballots was used to calibrate each scale, and both scales' weights and counts were recorded. The counting scales were then used to estimate batches of about 200 ballots apiece. (After some experimentation with different sizes, we decided that this size worked well for retabulation.)

Step 3. Joe Vitale of Election Systems & Services performed the retabulation itself. The DS850 was configured not to distinguish and "outstack" ballots containing overvotes and/or write-ins, so all ballots remained in their original order except in case of error. Vitale placed each batch of ballots in a mechanical "jogger" to minimize small differences in orientation, then scanned them. The imprinter printed (in blue ink)[37] a ten-digit serial ID number at the lower left corner. When ballots misfed and were outstacked, Vitale reoriented them and rescanned them at the end of the batch. From time to time, the scanner jammed and Vitale had to interrupt the scanning to clear the jam.

Step 4. The ballot manifest was sequentially prepared by Steve Taylor at this step, in consultation with staff at other steps. After retabulating each batch, three ballot counts were available: a scanner count and two scale counts. We also knew the first and last imprinted ID in each batch, which implied a fourth count (last ID minus first ID plus one). Any differences between this implied count and the scanner count were expected to be attributable to rescanned ballots. Indeed, each such difference coincided with the number of rescanned ballots. We considered the scanner counts to be definitive.

After each precinct was retabulated, we compared the sum of the batch counts to the counts reported on the Election Certificates (except for mail ballots) and to the original tabulation counts excluding provisional and manual-count ballots. All these counts matched, except that three ballots had been manually added to the results after the initial counts. Thus, we had high confidence in the batch counts and in the total ballot count, 9,021 voted ballots.

Each batch was put in a folder and each folder was assigned an ID that incorporated its precinct number, its box letter, and its serial position within the box. (This arrangement allows box letters to be reused, but because only ten boxes were needed, they were uniquely labeled A through J.) For instance, batch 0208-H-4 was a batch from precinct 0208, stored in box H, in the fourth folder.

Step 5. We experimented with several kinds of ballot boxes to store the batches: the metal containers presently used to store ballots; the plastic tote bins provided with the DS200 precinct-count scanners; generic plastic tote bins that were somewhat smaller than the DS200 bins; and cardboard boxes. The containers were evaluated on ease of retrieving specified batches, ease of moving, sealability, and ease of opening and closing. In Bristol, just one box per precinct was needed. The election officials at this station added folders as they were completed, then sealed the boxes with numbered seals and set them aside. These officials consulted with Steve Taylor to ensure that every batch ended up in the expected box.

37. In experiments, blue seemed easier to distinguish and read than either black or red.

Generating the Random Seed and Ballot Sample

The random seed used to select the ballot sample was generated by observers at the public audit, using a variant of the procedure used in Colorado and in many RLA pilots. Slips of paper bearing observers' names were placed in one hat; twenty ten-sided dice of different colors were placed in another hat. Observers took turns drawing name slips from the first hat. The observers whose names were drawn then took turns drawing and rolling a die, until 20 names had been drawn. The last three observers to roll dice rolled their dice simultaneously, to obviate any concern that the last person to roll could maliciously influence the audit result based on prior knowledge of the first 19 digits.

The 20-digit random seed was entered into the Audit Conductor software. The software then generated a simple random sample of 100 ballots from among all the Bristol ballots included in the audit. The ballots were identified by ballot identifiers (IDs) and also by their imprinted ID numbers. Each ballot ID consisted of a batch ID plus a sequential location within the batch, counting from the first ballot scanned. For instance, ballot 0203-C-1-29 would be the 29th ballot from the beginning of batch 0203-C-1. As the pilot design specified, the first 50 ballots in this sample (in random order) were designated as the "single-contest sample": only Question 2 would be audited. The remaining ballots were designated as the "ten-contest sample": all statewide and federal contests would be audited.

Retrieving the Ballots in the Sample

This process was relatively straightforward. All the voted ballots were divided among six stations, with a team of two election judges at each station. Each team received a software-generated pull sheet that listed, in sorted order, the ballots to retrieve by ballot ID and the actual sequentially imprinted ID. (Scanning reverses the order of the ballots: the first-scanned ballots end up on the bottom of the stack. The teams bore this in mind when searching for each imprinted ID.) Each ballot pulled was clipped together with a cover sheet that listed its imprinted ID number, to facilitate sorting. A place-holder piece of paper was put into the stack for each ballot pulled. Board staff then combined the ballots retrieved by each team, making sure that first the single-contest sample, followed by the ten-contest sample, was complete and that the ballots were in the correct order.

Interpreting the Ballots in the Sample

Three staff members and two election judges participated in interpreting the ballots. This process used the Audit Conductor software, which ran on a laptop computer attached to a projector, so all observers could watch the data entry process. For each ballot, the first staff member checked both the cover sheet and the imprinted ID against the prompt displayed by the software. This official read off the last four digits of the ID, then placed the ballot on a document projector so all observers could see the marks on the ballot. The two election judges read their interpretations of the votes, and the second staff member entered them into the software. At the same time, the third staff member entered the interpretations on a manual tally sheet, one row per ballot, to provide an independent record of the results.

Election judges were provided with an illustrated set of guidelines for interpreting possible votes.[38] Then the first staff member handed the ballot to the election judges, who reviewed the ballot itself while the second staff member read off the software's summary screen of the interpretations as entered. If the election judges agreed, the second official instructed the software to save the interpretations. (If not, the second official could change one or more interpretations.) Then the ballot was set aside, and everyone proceeded to the next ballot.

Examining the Results

The Audit Conductor software was designed to display summary results immediately after all 100 ballots were interpreted. Due to the minor software bug mentioned above, some calculations had to be rerun.

2. Ballot Polling

The basic steps in the ballot polling audit were similar to the ballot-level comparison audit:

- Divide the Portsmouth ballots into batches and prepare a ballot manifest accordingly. (Again, this step was done the week before the public pilot began.)
- Using the same random seed as for the ballot-level comparison audit, using Audit Conductor to select a random sample of 200 ballots.
- Retrieve the ballots in the sample, using one of four methods (see below).
- Interpret one contest (for governor) on the ballots in the sample and enter into Audit Conductor.
- Examine the results.

Batching the Ballots

We decided to rebatch the Portsmouth ballots for the ballot-polling audit the day after the Bristol retabulation" – January 9, one week before the public pilot. We also decided to use the DS850 as one source of per-batch ballot counts, to be compared with scale counts as well as the original reports. For purposes of the pilot, we believed that obtaining scanner counts would provide valuable additional information without unduly increasing the overall preparation time. (In a true RLA, this step should not be necessary provided that per-batch ballot counts that do not rely on the voting system are available. Also, in practice, it would make little sense to feed all the ballots through the scanner without retabulating and imprinting them for ballot-level comparison.) Thus, the workflow was similar to the retabulation – but somewhat faster on average because there was no need to save scanner data for each batch.

38. Rhode Island does not currently have an adjudication guide, nor regulations, for ballots that are manually tallied. For purposes of the audit, an adjudication guide similar to that used by the City of Denver, Colorado was produced. However, at no point during the audit was it consulted because there were no disagreements between the election judges.

We varied the batch size, by dividing the precincts into varying numbers of roughly equal batches, to learn how retrieval time was affected. Having found that the Tree and CS-20 scales consistently reported similar weights, we used just the (higher-capacity) Tree scale in the ballot polling audit, and used a stack of 300 voted ballots from Portsmouth to calibrate the scale. Not unexpectedly, the resulting estimated counts were closer on average to the scanner counts, but still fluctuated. The scanner counts for each precinct reconciled with the Election Certificates and original tabulation results, indicating a total of 7,966 voted ballots excluding provisional and manual-count ballots.

Selecting the Sample

For simplicity, the same public seed used to generate via the Audit Conductor software, the ballot-level comparison sample was used for ballot polling.[39] The ballot-polling sample consisted of 200 ballots in all, divided into four subsamples corresponding with four different retrieval methods. As in the ballot-level comparison audit, each ballot in the sample had a ballot ID that indicated its batch and its sequential location within the batch, counting from the top. As described below, the four retrieval methods used these ballot IDs somewhat differently.

Ballot-Polling Retrieval

The operational details of this pilot were quite complicated, because four retrieval methods were tested. In all four methods, whenever a ballot was retrieved for the sample, it was replaced by a placeholder sheet and attached to a cover sheet that identified it. Here is an overview of the process:

Scale method: First, 64 ballots were retrieved using one of three counting scales (one CS-20 scale and two Tree scales as described above). The precincts with the smallest batches were assigned to the CS-20 because of its smaller capacity; the remaining precincts were divided roughly equally. Each scale was calibrated against a stack of 300 voted Portsmouth ballots.[40] The staff at each station received a software-generated pull sheet that listed the ballots to retrieve, by batch and sequential location within the batch, counting from the top. Ballots in the first 10 or the last 10 of each batch were to be retrieved by counting, because this seemed easier for so few ballots compared to using a scale. Other ballots were to be retrieved by adding ballots to the scale until the desired count had been attained.[41] Although we knew that these counts could be off by a few ballots in either direction, minor innocent errors should not bias the sample in favor of any particular outcome or selection. (However, the possible extent and effect of retrieval errors has not been systematically studied.)

39. Internally, the Audit Conductor software generated 300 random numbers; it used the first 200 of these for the ballot polling sample, and the last 100 for ballot-level comparison.
40. As discussed later, a possible enhancement would be to recalibrate the scale for the reported number of ballots in each batch.
41. The audit teams were shown how to preserve the original order of the ballots by flipping ballots onto the scale, with the topmost ballot face down on the scale.

Next the ballots from the precincts were divided into roughly equal numbers of ballots and assigned to one of six stations. Again, staff at each station received pull sheets directing them which ballots to retrieve.

Counting method: Only eight ballots were retrieved by this method. (One team realized that it could retrieve ballots near the bottom of a stack by counting up from the bottom – a shortcut that could be formally supported if this method is used in future audits.)

Ruler method: A total of 64 ballots were retrieved by this method. Each team was provided a combination square: an adjustable "T" (or "L") with a millimeter scale. As with the counting method, ballots near the top or bottom of a batch were to be retrieved by counting. The remaining ballots were to be retrieved by setting the combination square to the appropriate measurement in millimeters (listed on the pull sheet), scraping the combination square across the top of the batch to cut (split) the batch, and then taking the top ballot below the cut point. As with the scale method, the ruler method may introduce small biases into the ballot sampling process.[42] Further research and experimentation is needed to understand the nature and consequences of such biases, and to suggest procedures for mitigating the effects of such (probably small) biases on the audit results.

k-cut method: A total of 64 ballots were retrieved by this method. The pull sheets told each team to retrieve a particular number of ballots from each specified batch. For each retrieval, teams were instructed to cut the designated batch six times, each time placing the ballots previously on top underneath. For each cut, the teams used an online random number generator to produce an integer between 1 and 99, and then attempted to cut the batch at approximately that percentage from the top. (For instance, if the number was 32, they would try to cut about one third of the way down, but a bit less.) After completing this process of six successive cuts, the ballot on top became the ballot in the sample.[43] Experimental and simulation results indicate that the k-cut method can generate approximately uniform samples, perhaps with far fewer than six cuts.[44]

Interpreting the Ballots in the Sample and Examining the Results

Ballot interpretation proceeded in essentially the same way as in the ballot-level comparison audit except that just one contest – for governor – was audited on all 200 ballots. Again, the summary results were displayed immediately after the interpretation step was finished. Because ballot polling audits do not allow the audit interpretation of each ballot to be compared to the machine interpretation, the audit could not produce specific discrepancies to investigate.

42. The ruler method may be more susceptible to sampling bias than the scale method, because the ballot retrieved may depend on how hard the audit judge presses the combination square downward. Also, we know less about possible variability in ballot *thickness* than about the observed variability in ballot *weight*.
43. This method was tested last because, unlike the other methods, it alters the order of ballots.
44. Mayuri Sridhar and Ronald L. Rivest, "k-Cut: A Simple Approximately-Uniform Method for Sampling Ballots in Post-Election Audits," *Cornell University*, November 21, 2018, https://arxiv.org/abs/1811.08811 (explores the method, justifies the choice of k=6 cuts, and suggests possible mitigations for any deviations from uniformity).

3. Batch Comparison

The batch comparison pilot tested variants of two common methods for manually counting ballots.

The *Sort-and-Stack Method* involves first sorting ballots into stacks for a contest by vote (or as “no valid vote”) and then counting each stack to determine the results.

The *Hashmarking Method* involves reviewing one ballot at a time while making hashmarks for each vote (or for “no valid vote”) in one or more contests, and then totaling the hashmarks for each selection.

The pilot implementations ensured that judges double-checked all vote interpretations and counts, that observers could verify all the work. These implementations prioritized accuracy: hasty methods that produce wrong results ultimately squander resources and confidence.

The sort-and-stack method employed two teams of two judges. Each team was given approximately one-half the ballots for the assigned polling place. First, the ballots were all oriented in the same direction, same side up. For each of the three contests counted: the two judges on each team sorted the ballots into stacks for each candidate (or no-vote). Then both judges reviewed that ballots were correctly stacked for each candidate in each race. Then both judges counted each stack. When counts did not match, stacks were recounted until the counters agreed.

The hashmarking method employed three teams of four judges each, later augmented by two more teams of four. They counted a polling place with about three times as many ballots as in the sort-and-stack method. First the ballots were all oriented in the same direction, same side up, separating out write-ins. The teams jointly counted the ballots into piles of 25, plus a pile of remaining ballots for both regular and write-in ballots. The ballots were then separated into stacks of 50 ballots plus extras. For each stack of 50 a team of four were given two preformatted hashmark sheets to record hashmarks for all candidates and no-votes in all contests. One judge read each vote on each ballot, while another official observed and checked the reader, and two officials made hashmarks on the two sheets (see Appendix E). At the end of each stack, the hashmarks on the two sheets were compared, and the contest totals were compared to the number of ballots in the stack. When discrepancies were found, the teams either recounted the batch with new hashmark sheets or counted the ballots for each candidate in question until they were satisfied that the result was correct.

The results for each sorting and stacking team, and for each of the hashmarking batches, were copied from input into a spreadsheet by a team of two, one doing data entry and the other checking the data entry. They were then compared to the tabulations and any discrepancies were investigated.

Rhode Island Pilot Timing Data

The RIRLA Working Group collected timing data during the pilot to compare the efficiency of the different approaches and the resources required to conduct each. Prior to the pilot, the Working Group developed forms to capture this data. (Copies of the timing forms are included in the Appendix.) After receiving a brief overview and training of the process, volunteers (election staff, computer scientists, students, etc.) helped collect the data by timing the various audit steps and recording it in the forms. The data collection process would have benefited from more clarity in the instructions and additional time to train the timers. Overall, however, the exercise went well and produced data that should be useful for Rhode Island as the state decides how to implement its audit law.

Below is a summary of the results, the data gathered in timing each phase. The state should consider the data, while understanding that they are specific to the pilot. Caution is warranted in generalizing them to future audits, particularly those that use substantially different procedures.

General

Rolling dice & recording random seed (used to determine the random samples):

- Rolling 20 ten-sided dice and recording the numbers took **14 minutes**.

1. Ballot-Level Comparison Audit

Rescanning & batching ballots subject to audit:

- Rescanning and batching the ballots took **approximately four hours**, which included some delays needed to discuss the process. The team processed **approximately 3,240 ballots/hour** (this excluded extended delays, but included minor ones, e.g. paper jams).

Median time to retrieve each selected ballot (by ballot type/location):

- All ballots: 45 seconds
- Ballots from a new box: 61 seconds
- Ballots from a new folder: 48 seconds
- Ballots in same folder and box: 31 seconds

Comparing each audited ballot to the initial result:

- The average time to evaluate one contest per ballot was **25 seconds**.
- The average time to evaluate ten contests per ballot was **62 seconds**.

2. Ballot-Polling Audit

Dividing ballots into batches & preparing ballot manifests:

- It took **approximately one hour and 40 minutes** to separate ballots into batches and prepare the manifests, an average rate of **4,770 ballots/hour**. (This included the time needed to rescan the ballots and to count the number of ballots per batch.)

Average time to retrieve and evaluate each selected ballot for each method used:

- The pilot collected data using two timing measurements: the overall time for each counting method and the time for each ballot individually. The overall time included discussion and time taken to "rework" the process. Due to a significant learning curve, the first few ballots typically took longer than the average.

	Average based on overall time	*Median based on individually-timed ballots*
• Scale method:	78 seconds	35 seconds
• Ruler Sample method:	78 seconds	75 seconds
• K-cut method with k=6:	104 seconds	86 seconds
• Manual count method:	230 seconds	104 seconds

As expected, the two timing measurements yielded different results. In general, however, they confirmed the relative speed of the different methods. Much of the difference between the mean and median can be attributed to "startup costs" (e.g. initial training and confusion). It also took into account the time spent opening ballot containers and, if necessary, retrieving them. Data were not collected for time required to position ballot containers at the audit retrieval stations because it was both rather brief and not expected to scale up to larger audits with different physical arrangements. (See the section "Considerations for Future Audits in Rhode Island" for further discussion about ballot container management.)

Notably, the batches in the pilot were relatively small. (The largest Portsmouth batch contained just 341 ballots.) By dividing the ballots into smaller batches, we were able to reduce the burden on the ballot retrieval teams and facilitate very careful ballot accounting (as the Election Certificate figures were checked against both scale counts and independent scanner counts). In a full-scale ballot-polling audit, a batch might consist of all the ballots cast at a precinct – perhaps in excess of 2,000 ballots. One should therefore exercise caution when using the data from this pilot to generalize what an audit conducted with substantially larger batches might yield (especially if conducted without pre-batching the ballots).

Comparing each audited ballot to the initial result:

- The average time to evaluate one contest per ballot was **25 seconds**.
- The average time to evaluate ten contests per ballot was **62 seconds**.

3. Batch Comparison Audit

Organizing ballots for audit (orienting & batching):

- Hashmark method: 6 seconds/ballot[45]

Interpreting and counting ballots:[46]

- Hashmark method: 7.2 seconds/ballot
- Sort-and-stack: 7 seconds/ballot per contest

45. The time varied substantially across groups, due to some start-up and training issues. The minimum average observed was 6 seconds per ballot; we believe that after training, other groups would achieve this rate.
46. The numbers for evaluating each method are not directly comparable. Sort-and-stack involved two persons orienting the ballots, sorting them by selection for **one contest** and then counting the stacks. The hashmarking involved two persons for orienting and batching the ballots and then teams of 4 for tallying **all the audited contests**.

Pilot Results

This section contains both the formal quantitative results of the pilot and a description of some of its immediate qualitative results.

1. Ballot-Level Comparison

Ballot Retrieval

Retrieving the sample of ballots for this method was fairly simple because it relied on the imprinted IDs. With 100 ballots divided across six teams, retrieval was completed quickly. However, a procedural error was detected and corrected during this stage. A retrieval team discovered that it had two pairs of cover/placeholder sheets that did not correspond to any of the ballots on its pull sheet. Within a few minutes, we determined that the discrepancy stemmed from an error when the ballot manifest was created, and that the ballots indeed belonged in the audit sample.[47]

Ballot Interpretation

Overall, it took approximately one hour and 15 minutes to audit a total of 550 vote opportunities (50 x 1 + 50 x 10). It took less than twice the time to audit ten contests on a ballot as it did to audit just a single contest on a ballot.

Note that there was a discrepancy among the 550 vote opportunities. One of the ballots that had been adjudicated in the audit as "Approve" was listed in the Cast Vote Record as "Reject." The ballot (and its scanner image) clearly shows a vote for "Reject." The error is most likely due to the audit team having inadvertently interpreted the vote for Question 1 instead of for Question 2 (a relatively easy mistake to make given the positioning of the questions on the ballot). Pilot observers failed to take notice of the error. This type of error can be reduced in a full-scale audit by having outside observers to monitor the audit team.

Despite this error, which resulted in a two-vote overstatement (the worst possible outcome in auditing a ballot), the measured risk for this and all other federal and statewide contests was, for the most part, well under 1% (or 0.01). Specifically, the risk was approximately 0.54% for the U.S. Senate race; 0.34% for U.S. Representative (District 1) race; 0.28% for Question 2; just above 0.01% for Secretary of State; and under 0.01% for the remaining contests. Thus, auditing just 50 or (for the U.S. Senate contest) 100 ballots produced very strong evidence that all the outcomes in Bristol would be confirmed by a full manual ballot count.

47. Specifically, we found that one batch's precinct was entered in the ballot manifest as 0025 instead of 0205. The random sample included two ballots from this batch, which did not appear on any pull sheet because "precinct 0025" was not assigned to any team. Ideally, the audit software would have detected this impossible precinct value at the start of the audit, or would have detected that some sample ballots did not appear on any pull sheet.

② Ballot Polling

Ballot Retrieval

As described earlier, the pilot used four different methods of ballot retrieval for the ballot-polling audit. Retrieval teams received training on each method at their stations immediately prior to this step of the audit. While conducting the training shortly before retrieval was useful for the pilot schedule and allowed observers to watch, it also curtailed the time available to teach the methods. This contributed to some confusion and minor errors during the actual retrieval, which were documented. The errors were minor and addressed as they arose. They did not threaten the integrity of the pilot.

Ordering Mistakes

Three of the methods require the ballots to remain in order (the k-cut method does not), but preserving the original order of the ballots was confusing. Because ballots neither bear numbers nor are oriented the same, teams should handle them carefully to preserve the original order. In reality, as long as there is assurance that the ballot is selected without the selection team seeing its contents, a mistake in the ballot order should not materially affect the audit results, but would make it harder for observers to confirm that the audit was performed correctly.

We observed various mistakes that altered ballot order, many of which could have been averted simply by placing a sticky note on the front of the topmost ballot. Systematic training on how to handle the ballots to preserve order is important. Presumably, it would be easier in an actual audit, where one would not have to learn four retrieval methods, but one.

The retrieved ballots posed a different kind of order problem. Teams were instructed to keep the samples for each method separate and in the same order as listed in the pull lists. However, the instructions were poorly communicated, and the ballot order was inconsistent. Ultimately, the 200 ballots had to be separated and sorted manually using the ballot IDs on the cover sheets. This process took several election officials about 20 minutes.

Scale Method

The scale method appears to have been the fastest one employed during the pilot. Presumably, it would have taken longer if the batches were larger (and higher-capacity scales would have been required). The teams also found the method straightforward and intuitive (and seemed to enjoy the adding and removing of ballots until the correct number had been selected).

Counting Method

On average, the counting method took appreciably longer than the other methods, even with the relatively small batches used in the pilot. (However, some teams eventually realized that, when retrieving a ballot near the bottom of a batch, they could save some time by counting from the bottom of the batch up.) Municipalities that use the counting method should consider adding pull sheet information that facilitates retrieval in this manner, and take into account the potential for confusion in their training.[48]

Ruler Method

The ruler method was faster than the counting but slower than the scale method. The initial retrievals took considerably longer than the successive ones, which could be attributed to the teams learning to use the combination squares (the "rulers"). While some teams adjusted quickly, others struggled to set their devices to the correct measurement. The difficulty was compounded by poor lighting and a lack of familiarity with both metric system scales and the rulers themselves. Most of these problems could be mitigated through training and better preparation, and by obtaining rulers with a better contrast.

The method does have some shortcomings that are difficult to overcome. First, it imprecise. Using an average figure to convert ballot number to millimeters both made it unlikely that the team would retrieve the "right" ballot and it increased the likelihood of including some ballots in the sample over others. This creates the possibility of bias. Also, it was time-consuming and awkward to use the ruler to select ballots towards the bottom of the pile. Further, it became more difficult to scrape the ruler across the top of the batch (to select a ballot) as the batch size and weight increased. For large batches, one might select a ballot by sticking a sharp object (e.g. a nail or flathead screwdriver) at the base of the ruler where the scrape would occur. (And even this would be difficult to execute if 1,000 ballots were stacked above the ruler.)

k-cut Method

The k-cut method proceeded at about the same rate as the ruler method, but it may be easier to use for large (but physically manageable)[49] batches. It can also be conducted without special equipment. And, as previously noted, it may be possible to obtain approximately uniform samples with fewer than six cuts, which could speed up the process considerably.

While some teams liked using this method, others found it difficult or uncomfortable to make the estimates it requires (e.g. in a stack, which ballot was 37% of the way down). However, this concern might be allayed with training that clearly explains the rationale for repeated cuts.

48. For instance, ballot number 572 in a batch of 703 ballots could be annotated on the pull sheet as "132 from the bottom."
49. What "physically manageable" means depend on the size of the ballots and whether there is one person making the cuts, or two people working as a team from different sides of the table. With two people, a stack of 1500 ballots may be manageable.

Ballot Interpretation

It took approximately 90 minutes to audit the 200 ballots for the governor's contest. The audited sample contained 99 votes for Gina Raimondo, the reported winner; 69 votes for Allan Fung, the reported runner-up; 23 votes for other candidates; and 9 undervotes. The sample margin, 15.7% of valid votes [(99 - 69) / (99 + 69 + 23)], was somewhat smaller than Portsmouth's reported margin of 22.2%. The measured risk for this audit is 10.2%. If it had enforced a 10% risk limit, the audit would have had to continue (although probably not for very long).

3. Batch Comparison

Both methods, hashmark and sort-and-stack, were designed to be self-correcting. In a few batches the two hashmarkers' counts did not match, in those cases, the teams had to recount those batches. Likewise, a sort-and-stack team had to redo the process for one contest because, while re-checking the stacks, they realized that they had mixed up stacks between candidates.

Both methods produced results that matched the original results exactly, proving that the methods, officials, and supervision were all effective. Had they not matched, depending on the level of differences, the data entered would have been rechecked and, if necessary, the counting repeated until the manual count results were accurate. Note that accurate manual count results can differ from scanner counts due to non-standard voter marks.

50. A 2x2 chi-square test shows no statistically significant difference between the two leaders' audit vote counts and their reported total (Portsmouth) vote counts ($p = 0.37$). Thus, the difference in margins between audit and reported totals is well within the range of expected sampling variability.

Lessons from the Pilot

Here we briefly consider some lessons from the pilot that are worth considering in future pilots and audits. Some of these could apply to any audit; others are method-specific.

General Lessons

Think hard about how to combine transparency and efficiency. Some election officials conducting RLA pilots have decided to avoid advance publicity and to invite few or no outside observers, so they could focus on learning and experimenting without the distraction of explaining the work (and any complications or mistakes) to an audience. The Rhode Island pilot took a very different approach, inviting and welcoming dozens of outside observers, and experimenting with unfamiliar methods knowing that some were bound to work better than others. This approach offers real rewards: it brought considerable public attention to the Board of Elections' efforts and helped local election officials and other observers understand the logistics of post-election audits. But it also has real costs, both tangible and intangible. We did decide to complete some processes, such as the retabulation, prior to the public event so that the workers could be narrowly focused on the task. At the same time, we wanted observers to be fully apprised of these processes. (A short video documented the retabulation.)

One tangible cost of having observers was the time spent setting up the space for observation, including projectors to display the ballots and the audit software, as well as extensive seating. This paid staff time was well spent: observers appreciated being able to watch the ballot interpretation process without having to crowd around the officials – and simply having a comfortable place to sit – and this arrangement also was more comfortable for the officials.

We chose to train the ballot retrieval teams in the various methods during the event, with observers present. In retrospect, it likely would have been better to train the teams in a calmer environment where they could master the methods without the pressure of having many unfamiliar people watch. The team members did not complain, but surely they faced an uncomfortable situation. Also, the trainers (including Board staff) had to balance answering observers' questions with being available to help the retrieval teams.

It would have been helpful – and would be very important in an official RLA – to have written explanations of the various procedures. The more information that can be provided to observers in advance, the easier it is to manage their questions during the procedures themselves.

Don't skimp on staff. As with any election procedure, adequate staffing can make the difference between a manageable task and a morale-crushing slog. This need posed a hard choice for the pilot. The Board of Elections actively considered involving volunteers -- potentially including members of local Boards of Canvassers -- in the pilot processes. In retrospect, this could have been valuable because the local boards may play a larger role in future audits. However, the pilot involved complex staffing needs that could change throughout the day, and the Board wanted to ensure that no portion of the pilot would be slowed or impacted by lack of staffing. Therefore, it opted to rely more on bringing back paid election seasonal staff, instead of relying on volunteers. This choice worked well in that the seasonal staff did excellent work. In retrospect, many members of local Boards of Canvassers attended as observers, and could have been incorporated in the pilot. That said, feedback from the Boards of Canvassers indicates they appreciated being able to observe all three pilots.

Test the workflow and ergonomics. At several points in the pilot, we either adapted our arrangements based on immediate experience, or endured the (minor) costs of our failure to experiment in advance.

During the retabulation the workflow (see page 24) was rearranged when a more efficient order was discovered. This led to the creation of the batch tracking sheet that was incorporated into the process.

Lighting conditions provide an example of enduring rather than adapting. For most purposes, the basement warehouse space would be considered well lit. However, for the fine detail work of retrieving and examining ballots, the ambient light was sometimes inadequate to avoid real strain. If we had experimented with these processes in the spaces used in the pilot, we probably would have brought in task lighting to ease the burden on workers.

What we learned about containers. As mentioned before, the pilot tested four kinds of containers to store voted ballots after dividing them into batches: the metal containers currently used to store voted ballots, and three alternatives. The containers were qualitatively assessed based on the following criteria:

- Ease of access to individual folders
- Ease of moving
- Sealability
- Ease of opening and closing

All these containers performed reasonably well. The DS200 ballot bins, which were designed to hold large numbers of ballots, were deemed somewhat unwieldy because of their size and weight. The generic plastic bins were more manageable than the DS200 bins and were easy to seal (using two plastic wire seals) and to open and close. When they were reasonably full, they allowed the ballot folders to stand up, making it easier to retrieve individual folders. (This advantage did not matter much in the pilot because most batches were represented in the ballot samples.) The metal and cardboard ballot containers held fewer ballots than the plastic bins -- which also simplified retrieving individual folders and limited the maximum weight of these containers. The cardboard containers were lighter than the metal containers, and sealed well, but the flaps made them difficult to open and close. Board staff felt that, on balance, the existing metal containers were most practical for their purposes.

Plan ahead for software acquisition, development and testing. Robust fully featured software for all the planned auditing methods was not available leading up to the audit, so as described above, we devoted considerable effort to developing new software which met the needs of the pilot. The landscape of auditing tools is rapidly changing, and advanced planning remains essential, both for pilots and for the more rigorous software and logistical requirements of actual audits.

Method-Specific Lessons

Retabulation: batch size and management matter. The ballot-level comparison audit, as noted earlier, required ballots to be rescanned using the DS850s. Through trial and error it was determined that batches of approximately 200 ballots worked best. Batches of 300 created appreciably more work when a misfeed occurred; batches of 100 were so small as to be inefficient. Also, because of the frequency of misfeeds, it is important to wait until after the entire batch has been scanned to enter the numbers into the ballot manifest. We found it helpful to create a tracking sheet placed on top of each batch. These tracking sheets were used to record timing and count information – and to ensure that the batches stayed in precinct order.

Retrieval using the ruler method: Some teams found that the combination squares used during the pilot worked best when the ballot stacks were placed at the edge of the table. Also, the combination squares used in the pilot were acceptable, but squares with high-contrast (e.g., white-on-black) rulers would be preferable.

Retrieval using the scale method: During the retabulation, we found that the scales consistently undercounted, most likely because the ballots used to calibrate the scanners were test ballots that were slightly heavier on average than the voted ballots. We also observed that the average ballot weight varied slightly but appreciably from batch to batch, generally falling within about 0.5% of the overall average. We believe that most of this variability occurs when the ballot stock is manufactured; storage conditions also can affect ballot weight. Because of this variability, scale counts can vary by plus-or-minus one for every 200 (or so) ballots, and possibly more.

With this in mind, a protocol enhancement worth considering for the scale method would be to recalibrate the counting scale for each batch based on the number of ballots in the batch, as reported in the ballot manifest. This recalibration simply entails placing the entire batch on the scale,[51] entering the number of ballots, and hitting one more key to enter it as a count, which should take only a few seconds per batch. Recalibration compensates for differences in ballot weight from batch to batch.

51. At this point, audit judges can compare the scale's count of ballots in the batch against the figure in the ballot manifest. These generally should be similar within a fraction of a percentage point.

Comparison with the "Principles and Best Practices"

An *ad hoc* group consisting of former election officials, election security advocates, and scholars with expertise in relevant fields developed a set of guidelines for designing and implementing high-quality post-election audits. These guidelines were published in 2018 as the *Principles and Best Practices for Post-Election Tabulation Audits ("Principles and Best Practices")*, an update to a similarly-named document published in 2008. Comparing the pilot audit design and conduct to the nine general principles described in the *Principles and Best Practices* not only provides some basis for gauging its success, it also offers critical insight into how the audit can be improved.

Overall the pilot's conformity with the *Principles and Best Practices* was very good. However, because this was a pilot project that covered a relatively small number of ballots from just a few local jurisdictions, it was impracticable to comply fully with some of the principles.The points below not only explain when this was the case, but they also note where the Rhode Island audit law contemplates conduct that is different from what was feasible during the pilot.

Examination of Voter-Verifiable Paper Ballots: The pilot adhered to this principle. Paper ballots marked by hand or by AutoMARK were used throughout the pilot, and the voter ballot marks (i.e. not barcodes) were the basis for tabulation. Rhode Island will continue to use voter-verifiable paper ballots.

Transparency: With the exception of the ballot re-scan, and publication of the data to be audited, the pilot was open to the general public, and adhered to best practices of observability. It permitted those in attendance to view the ballot retrieval and adjudication from a short distance. (This was adhered to despite the crowding that occurred at certain points during the pilot.) The room setup included two large screens to make it easier to view critical audit steps such as the dice roll to create the random seed and the ballot adjudication. The ballot-level comparison results could have been presented in a more accessible format or some manner that would have made it easy for the general public to understand them without detailed explanation. While the pilot result data were not published, this could have been done, and state law requires the publication of future audit results within two days of the Board's acceptance of them. (R.I. Gen. Laws § 17-19-37.4.)

Separation of Responsibilities: State election officials at the Board had complete discretion over the conduct of the pilot, which technically violates this principle. However, the pilot was designed and conducted in partnership with outside groups and individuals who possess expertise in election administration, election security, and post-election audits (though without input from the general public). According to state law, the Board, which is responsible for tabulating state election returns (R.I. Gen. Laws § 17-19-37), has the authority to develop audit regulations and policies, including those related to selecting the contests and number of units to be audited. (R.I. Gen. Laws § 17-19-37.4.) Those regulations

52. The document is available at https://electionaudits.org/principles.

and policies should be developed after taking into account input from the public and other election officials and stakeholders (e.g. the Office of the Secretary of State, Boards of Canvassers, etc.). They should also provide for segregation of duties, as well as recusals when individuals conducting the audit have close ties to, or are themselves, candidates for office.

Ballot Protection: Rhode Island law provides for sufficient ballot protection from the time of the vote through their delivery and storage at the Board. (R.I. Gen. Laws § 17-19-19, 33, 35 & 39.1.) While the pilot did not include an observation of the post-election chain-of-custody process, there was some inquiry into details of ballot storage during pilot development. The ballot containers were sealed at the beginning of the audit; Board staff was present during the retabulation of the ballots; a thorough ballot reconciliation was conducted as ballots were rebatched (the reconciliation reports should be part of the record); and the ballots were returned and the containers re-sealed at the conclusion. In addition, ballot anonymity was preserved. In general, the chain-of-custody procedures appeared sound, and the pilot adhered to this principle.

Comprehensiveness: Provisional and "manual count" ballots were disregarded for the purposes of the pilot. Otherwise, all voted ballots from the three participating jurisdictions were subject to being audited. Also for the purposes of the pilot, audited contests were predetermined, prioritizing federal and statewide contests.

Appropriate Statistical Design: The statistical design was appropriate for the audit. Each method was designed to use as few ballots as possible to provide evidence that the reported outcomes were correct, while still offering people a meaningful opportunity to participate, observe, and learn in the experiments. With regard to the design, there are two notable points. First, the "reported outcome" did not apply to the statewide results but the results in the jurisdiction being audited. This was circumscribed appropriately to account for the samples being used. Second, for the ballot polling audit, the measured risk was 10.22%, which was slightly above the aspirational 10% limit. However, the pilot was designed to audit a predetermined number of ballots. If this had been a full RLA that enforced a 10% risk limit, the audit would have continued in accordance with state law (which calls for an escalation in ballot counting until the audit results show "strong statistical evidence" that the outcome was reported correctly). (R.I. Gen. Laws § 17-19-37.4.)

Responsiveness to Particular Circumstances: The pilot limited the number of ballots to be audited and, by design, did not contemplate or provide for additional auditing. The law provides the Board of Election broad authority, including to draft regulations that provide for targeted sample auditing of ballots, precincts, or offices in the event that officials suspect voting irregularities or circumstances otherwise warrant increased scrutiny.

Binding on Official Outcomes: The pilot audit was conducted in January 2019, two months after the 2018 election results were certified. Therefore, it was not feasible for the pilot to meet this goal (nor was it contemplated that it would). According to state law, however, election officials must "audit [the] tabulation of the unofficial final results," including recount results, prior to certification by the Board. (R.I. Gen. Laws § 17-19-37.4.) The law contemplates that the results of future audits will be binding on the official outcome. Future audits will be designed to adhere to this principle.

Investigating Discrepancies and Promoting Continuous Improvement: The pilot methods worked well for addressing the anomalies that were observed during the pilot. As explained above, the one outright discrepancy found (in the ballot-level comparison) demonstrably was caused by an error during the audit itself. However, the same procedures could have quickly isolated errors in the original machine count regardless of their cause(s).

Part IV: Considerations for Future Audits in Rhode Island

Rhode Island Audit Scenarios

Rhode Island must begin preparing to audit two elections in 2020: the April presidential preference primary ("primary")[53] and the general election in November. Below is an examination of the potential RLA methods[54] and related procedural options that Rhode Island may adopt. The state's decision among the alternatives will influence the scope of the audit (i.e. how many contests are audited and how thoroughly), the time and resources that will be required, and the level of transparency and public observability.

Assumptions & Considerations for All RLA Methods

Voter Turnout. In 2016, approximately 184,000 people voted in Rhode Island's primary election and 464,000 in its general election. Approximately 41,000 of the general election ballots were cast by mail. In 2020, we should expect about 500,000 people to cast their ballots in person for the general election and, according to the state's projection, about 40,000 to vote by mail.[55] For the purposes of this analysis, we focus on the general election. Given that fewer ballots are cast in the primary, it will serve as a "trial run" for conducting the General Election RLA in November.[56]

Contests. The state's RLA statute does not specify what contests must be audited. However, the National Academies of Science, Engineering, and Medicine consensus report recommends that RLAs "should be conducted for all federal and state election contests, and for local contests where feasible."[57] Accordingly, our analysis assumes that, at a minimum, federal and statewide contests will be subject to audit. However, the reader should note that there will be no statewide offices on the 2020 ballot, but there may be statewide measures.

53. Notably, Rhode Island will hold a primary election in April 2020, and a primary contest for state offices later in 2020. Here, the word "primary" refers to the former, which, under the RLA law, must be audited.
54. It is possible for the state to adopt a "hybrid" RLA, with precincts and jurisdictions employing different audit methods (which could potentially reduce the number of ballots to be audited or provide officials with more information). However, these hybrid approaches will not be discussed here.
55. Board of Elections estimates.
56. For the Presidential Preference Primary the Board of Elections must decide whether to audit based on the margin of victory or the allocation of delegates. No prior RLA has been conducted on allocation of delegates.
57. *Securing the Vote: Protecting American Democracy,* 9.

Risk Limits. Nor does the state's RLA statute establish a risk limit to which the races must be audited. The analyses below show estimated number of ballots to be sampled and estimated work loads for both 5% and 10% risk limits for all three methods.

Contest Margins. In past Rhode Island general elections, the outcomes of federal contests have resulted in moderately large vote margins. In 2016, for example, the Democratic nominee for president won the state's vote by 15.5 percentage points. David Cicilline, currently a representative, won his first election in the 1st congressional district by about six percentage points in 2010, and his margins of victory have increased since.

There have been similarly large margins for most recent statewide measures. In 2014, Question 3 (on whether to hold a constitutional convention) was rejected by approximately 10.2 points. That was the closest margin for a statewide measure since 2010. However, some statewide offices have had narrower margins. For example, the 2010 governor's race was decided by about 2.5 points.

For the purposes of this analysis, we assume vote margins of 15, 10, 5, and 2 percent, realizing that they will vary from contest to contest.

Audit Location(s). Rhode Island's Board of Elections needs to choose not only one of three alternative RLA methodologies, but also whether to conduct its 2020 risk-limiting audits in a central location or dispersed among a few or all of the 39 individual municipalities that are responsible for carrying out in-person voting in their individual precincts. Any of the three major RLA methods (ballot-level comparison, ballot polling, or batch comparison) could be carried out centrally because all but one of the 39 municipalities lie within an hour's drive of Providence. The ballot-level comparison method would most likely need to be carried out in a central location because all ballots from precincts would need to be transported to Providence and rescanned, tabulated and imprinted for a transitive ballot-level comparison audit. But either the ballot polling method or the batch comparison method could be dispersed over all 39 different municipalities or consolidated into a smaller number of distributed locations.

In general, one principal advantage of conducting risk-limiting audits at a single geographic location is that it is much easier to manage a centralized audit. As we saw during the pilot audits event in January, when problems arise, a knowledgeable person can be available relatively quickly in person to help solve them. It also is much easier for managers to notice when something is going wrong in some part of the overall process and deal with it quickly. Software only needs to run at one location, and software would not need to be supported in up to 39 different locations with quite different levels of computer expertise.

It should also be noted that the Board of Elections may be required to conduct a full retabulation of all the ballots to conduct a recount in a close statewide contest, which also would require transporting them to the Board of Elections warehouse and storing them there temporarily.[58] Although this has not occurred since Rhode Island adopted the use of hand-marked paper ballots and optical scanners, it could happen in any year that statewide elections take place. Doing regular risk-limiting audits centrally following each election would provide both capacity and experience that could make conducting full tabulation recounts less traumatic for Board and local staff as well as the general public.

Carrying out an audit in geographically dispersed locations has several advantages, including parallel operations and less impact on the central warehouse facility. But it also has a number of disadvantages, including the following:

- Proper training and management of the auditing staff in all 39 municipalities would be required. This management includes recruiting the auditors, training them, and running the audit. The documentation and training would have to be sufficient to make sure that the procedures were uniform across jurisdictions -- e.g., how to organize ballot boxes for easy retrieval, how to find sampled ballots, and how to enter data from sampled ballots into audit software.

- Each municipality would have to have a secure space for ballot storage and for the examination of the ballots, including sufficient space for observers.

- If municipalities consolidate to a smaller number of locations for the audit, they might have to find larger spaces, track which ballots came from where, and deal with other logistical issues.

- There would have to be coordination between all the municipalities to know how many ballots to audit in each municipality, to capture and analyze the results, and to determine when and how to escalate the audit if needed. Some of this could be accomplished through software, but the software would have to be set up in all audit locations (up to 39 municipalities) as well as at the central site, and all those software components would need to be coordinated from a central site.

- During the audit there would need to be help available via telephone and possibly even individual staff dispatched to help address problems that can arise in remote offices.

The Board of Elections will need to carefully weigh the issues, constraints, advantages and disadvantages of centralization versus decentralization as well as of the three different RLA methods. The Board needs to decide these questions as soon as practicable so rules and regulations can be drafted and another pilot audit can be conducted to test software and procedures for whatever method is chosen well in advance of 2020.

Items to freeze and publish before rolling of dice and sample selection. It is essential that certain materials be frozen prior to the beginning of the process for selecting the ballots to

58. The recount procedure is detailed in the *Board's Guide to Election Recounts*, available at http://www.elections.state.ri.us/documents/Guide_to_Recounts.pdf.

be audited. For all three methods, the ballot manifest should be frozen and published for public inspection prior to the rolling of the dice. For ballot-level comparison audits the cast vote records should similarly be frozen and published prior to the dice roll. Similarly, for ballot polling and batch comparison the precinct totals and subtotals should be frozen and published prior to the dice roll.

Alternative Risk-Limiting Audit Methods

The following sections discuss Rhode Island-specific implementation issues for each of the three major RLA methods, followed by a summary comparison of the estimated number of ballots that might have to be examined manually for each method given different contest margins and risk limits, along with estimated workloads and additional costs that would be required for each different alternative.

1. Alternative 1: Ballot-Level Comparison Method

Retabulation. Given Rhode Island's current voting system, a ballot-level comparison transitive audit would require retabulating and imprinting ballots from all precincts to obtain cast vote records matched to individual ballots. The ballots would also be tallied to verify whether the outcomes initially announced are consistent with the results of retabulation.

Mail ballots, which are an increasing portion of all ballots cast in Rhode Island (see above), can be imprinted during their initial tally. Proceeding in this manner for mail ballots would add a minimal amount of time to the staff work. (It would also permit officials to audit mail ballots through the ballot-level comparison method even if other methods are used to audit in-person ballots.)

In the days leading up to the pilot, the ballots were retabulated on an ES&S DS850 scanner. They were scanned at an effective rate of about 4,000/hour.[59] At this rate, assuming that an approximate 400,000 ballots will be cast in person for the 2020 general election, the state could expect to complete the retabulation in perhaps 170 to 190 scanner-hours, depending on the prevalence of two-sheet ballots. It is difficult to estimate the labor costs associated with transporting ballots to a central location and re-organizing them as part of the retabulation. However, these costs could be minimized by giving careful consideration to how ballots are initially organized and stored to facilitate more efficient ballot retrieval. And transportation from each precinct is already required for the DS200s.

Currently, the Board of Elections has just two DS850 scanners.[60] To complete the retabulation in a timely manner, there are several possible choices, including the following:

- Extend the time for retabulating the ballots. Using existing equipment the Board of Elections may be able to scan and imprint all ballots in a statewide race in a little over a week's time running running 12 hours per day. Either by starting the retabulation earlier, or by beginning the adjudication of ballots later, or both, the Board of Elections could retabulate all the ballots without acquiring additional equipment.

59. The DS850 is rated at up to 300 pages/minute, so the scanning could potentially proceed at a faster rate than it did during the pilot. However, the DS850 tends to jam frequently and has to be tended to with the utmost care. Also most of the scanner operator's time was not spent actually scanning the ballots, but handling them before and after the scan. Note that the scanner throughput of 4,000 sheets/hour is faster than the timing reported under "Rhode Island Pilot Timing Data" above, which includes tasks before scanning started and after scanning concluded.
60. Using the two available DS850 scanners and working from 7:00 am to 11:00 pm in shifts, each machine would process 64,000 per day (16 hours * 4,000 per hour), so 750,000 ballot sheets could be scanned in about 11.7 machine days, or 6 calendar days.

- Lease additional DS450 or DS850 scanners. Notably, during the pilot at least six election officials performed tasks necessary for retabulation, including opening, orienting, and batching ballots, creating the ballot manifest, and tasks related to ballot storage. These officials had considerable slack time, which could be used for scanning. We have heard that although DS450 scanners are not as fast as DS850s, they may not jam as frequently, so their throughput may be comparable.

- Consider using commercial scanners. Using a slower commercial scanner might increase the scanner-hours needed for retabulation by perhaps 25-50%.[61] Commercial scanners have imaging and imprinting capabilities. However, they do not have the capacity to perform tabulations to verify the results initially announced. To perform this step, the Board would have to acquire retabulation software from a private vendor or an open-source system.

- Issue an RFP for a short-term contract to retabulate and imprint an estimated 750,000 ballot sheets for the 2020 general election. (Potential bidders could include ES&S, Clear Ballot, and possibly other vendors who have capable hardware, software, and personnel.) If done, this should occur under strict supervision and management by Rhode Island Board of Elections staff, subject to public notice and observation, at its new facility in Providence.

Retabulation and imprinting demand a considerable amount of up-front work. However, it is an essential step for conducting a transitive ballot-level comparison audit, which is the most efficient and effective of the three RLA methods for races with close margins. It can also help isolate the cause(s) of any discrepancies. It also can support auditing of multiple contests – including “opportunistic auditing” of contests for which no risk limit is set – with minimal added work.

Ballot Samples. One can expect that auditing a contest with a 5-point margin to a 5% risk limit (with some tolerance for minor discrepancies) would require sampling just 150 ballots, assuming few discrepancies. The sampling plan in the table below can be adjusted to account for unofficial margins of the various contests, which would allow multiple contests to be audited. Depending on the number of statewide measures, it is likely that all statewide measures could be audited on 200 ballots and performed within three hours and 30 minutes (not including break time). We suggest increasing the minimum sample size of ballots statewide to provide good coverage of both U.S. House contests. Although we have not systematically studied down-ticket contests, it should be feasible to conduct opportunistic auditing of local contests on these ballots.

Ballot Retrieval & Audit Interpretation. During the pilot, it took an average of 68 seconds to retrieve the first ballot in a new ballot box (already positioned next to the table). While teams retrieved several ballots from each box during the pilot, in a live audit, it would be more common to retrieve just one ballot per box. Additional time would be required to retrieve and replace each box of ballots.

61. For example, the Fujitsu fi-6800, rated at 130 pages/minute, would operate at about half the speed of a DS850 scanner. In the pilot, however, the DS850 was operating less than one-third of the time due to tasks and delays unrelated to the scanner. The Fujitsu fi-6800 scanner costs about $19,000 and has an imprinter installed.

During the pilot, entering audit interpretations for one contest proceeded at about 140 ballots/hour (25 seconds/ballot). Entering audit interpretations for multiple contests was more efficient. It took just under 60 ballots/hour (62 seconds/ballot) to audit ten separate contests on a ballot in which the voter could make one selection among the choices. Nevertheless, it would be hard to sustain these rates of entry over long periods of time. When auditing lots of ballots, it would be helpful to allow multiple teams to simultaneously enter audit interpretations. Doing so can make observation more challenging, so this sort of parallelization should be discussed with interested parties.

See Table 2B for more details on estimated times and costs for different methods.

2. Alternative 2: Ballot Polling

Ballot polling audits can be conducted for multiple contests at once, including all statewide and federal contests. Unlike the ballot-level comparison method, ballot polling audits do not require ballots to be retabulated. This method is ill-suited for opportunistic auditing of local contests using the statewide sample, because it is unlikely to provide useful information given the small sizes of most statewide samples. Therefore, specific local contests would need to be audited separately, if at all.

Ballot Samples. The size of the ballot sample needed to complete a ballot-polling audit increases rapidly for narrow-margin contests (< 2%), and is unpredictable even when the reported results are accurate. Officials should therefore be conservative in planning for and setting an initial sample size, aiming for one that provides a large chance of attaining the risk limit without additional auditing (assuming that the reported results are accurate). Not only will a larger ballot sample reduce the likelihood of having to retrieve and audit additional ballots, it will also provide stronger evidence about the accuracy of the audited contests. The optimal sample size will depend on the reported margins and the relative ease of auditing additional ballots.

Local vs. Central Location(s). Ballot-polling audits can be conducted locally or centrally (and the results of audits performed either way can be combined). By conducting audits locally, officials avoid the need to transport ballots, store them in a central location, and them return them to local storage. Performing audits locally may also expedite audits by allowing many teams to work at the same time. It would have mixed effects on observability, requiring more observers, but also allowing observers to stay closer to home. Conducting audits centrally may simplify necessary training and make it more efficient to manage and communicate during the process. Central management would eliminate the need to combine ballot manifests and audit results from around the state, and to perform additional auditing in several locations. Central audits may also be more conducive to observation of the entire audit by election officials and members of the public.

Ballot Counting & Retrieval and Audit Interpretation. Officials can use Election Certificates and other precinct records to create the ballot manifests for the ballot polling method. (Ballot manifests can be prepared centrally or locally, and by several election officials working in parallel.) In preparing ballot manifests, officials ensure the accuracy of ballot counts. Among the approaches tested thus far, the fastest method of getting an accurate ballot count is to use counting scales. Once calibrated, a counting scale can check the number of ballots in a batch in well under one minute. However, using a scale for the ballot count can produce results that vary from a true count, perhaps by a fraction of a percentage point per batch. Any large discrepancies between reported counts and scale counts should be investigated and resolved.

During the pilot, the counting-scale retrieval method took a median 40 seconds/ballot, not accounting for retrieval of the boxes. Calibrating the scale to the number of ballots in each batch would have added a few more seconds per batch. This suggests that the overall retrieval rate per team would be approximately 105 seconds/ballot (or about 35/team/hour), which is similar to the estimate for retrieving imprinted ballots for the ballot-level comparison method. The rate would improve considerably for larger samples because it takes less than a minute to retrieve an additional ballot from an open box. In general, it should be possible to complete the ballot retrieval in a single work day. Note that this analysis may also apply to other retrieval methods. (For example, the k-cut method of retrieval could take more or less time depending on the number of cuts.)

If the three federal contests in 2020 were the only ones audited, one could expect to enter a full interpretation for 110 ballots/hour. If there were seven statewide questions, as there were in both 2012 and 2016, the process might slow to around 55-60 ballots/hour (which is similar to the rate for entering interpretations for the ten contests during the ballot-level comparison).

Carrying out a ballot-polling audit in a number of distributed locations could be challenging to manage. More training would probably be required to make sure that staff at each location are following the same procedures in terms of selecting sample ballots, marking their locations in their respective ballot boxes, entering votes from sampled contests into the audit software, and so on.

3. Alternative 3: Batch Comparison

Of the three RLA methods, the batch comparison approach requires sampling the greatest number of ballots. For purposes of simplicity, we assume the batch comparison method is used in an unusually close statewide contest, and the number of in-person ballots exceeds 1000 on average coming from more than 100 sampled precincts around the state. (Mail ballots could be audited using the ballot-level comparison method, and the results incorporated with the results from the batch comparison method to yield overall risk limits.) Gauging from the pilot, auditing a close statewide contest using sort-and-stack method should require an estimated 7 seconds/ballot using two-person teams for each race, plus an estimated 7 seconds/per ballot for orienting the ballots before counting the races. Using the more conservative pilot estimate, the batch comparison method (or full manual count) could audit about 510 ballots/team-hour. Auditing 100,000 ballots for one contest would take approximately 195 team-hours. Of course this can be performed by several teams auditing ballots simultaneously. The pilot results also indicate that the hashmark method could audit multiple contests on a subsample of the ballots at approximately 4 seconds/contest (and that there may be efficiencies gained as the number of contests being audited is increased).

Storage Considerations for Batch Comparison Audits. One attractive feature of a batch comparison audit is that sample batches could be selected (either precincts or machines) at the beginning of the audit and then ballot boxes for those batches could be transported to and stored temporarily at the Board of Elections warehouse for the audit. The total number of boxes for sampled batches would be considerably less than the grand total of 1,200 boxes that would need to be at the Board of Elections for a centralized audit using either ballot-level comparison or ballot polling methods.

Treating each DS200 as a separate batch would provide smaller batches to audit and thus fewer ballots to count.[62] Using scanners rather than precincts as the audit unit could reduce the batch comparison workload by close to 30% for margins from 2% to 5%. Basically, the number of batches in the sample would be roughly the same whether the batches were based on scanners or precincts, and the number of ballots in each scanner batch size is less than or equal to the number of ballots in the entire precinct. For instance, the 2014 margin for governor was about 4.5 points. To audit a similar contest by batch comparison, at a 10% risk limit, might entail counting about 63,000 votes at the precinct level, or 44,000 votes at the scanner level, at 2016 turnout levels.

62. Board of Elections staff have said that they fear in precincts that have more than one DS200, many election workers would not be able to be sure to put ballots from separate machines in separate boxes. However, other states, such as Maryland, have successfully been able to keep ballots from different machines separated in different boxes.

Audit Sample Sizes for Different Methods and Contest Margins

Table 1 presents estimates of the number of ballots one would need to audit in an election using each of the three major types of RLA methods. It shows the estimated number of ballots needed at four different margins, assuming that the reported counts are generally accurate. For batch comparison, the ballot estimates are based on data from the 2016 Rhode Island general election. (The number of batches depends only on the proportional margin.) The actual number could vary in other jurisdictions or other election years. These estimates use a 10% risk limit; to reach a 5% risk limit generally would require auditing about 30-40% more ballots.

For sake of simplicity, the estimates also assume a contest that requires voters to select between one of two candidates, and that there were no invalid votes cast. They also assume ballots consist of just a single sheet of paper. For audits with two-sheet ballots, one might need to double the number of sheets sampled.

As we mentioned earlier, ballot polling results are substantially less predictable than comparison results: additional sampling may be needed even when the original count is generally accurate. Here we present four possible ballot polling sample sizes for each margin. The smallest sample size has a 50% chance of attaining the risk limit, thus completing the audit, if the original count is accurate. Auditors might prefer to start with this relatively small sample size if it is not very difficult to expand the sample if necessary. Larger ballot polling samples have a larger chance of completion; the largest sample sizes shown here provide a 95% chance of completion.

Table 1: Estimated Number of Required Ballots to be Sampled and Audited, 10% Risk Limit
(by Auditing Method and Contest Margin)

		Margin of Contest Decision			
	Auditing Method	**15%**	**10%**	**5%**	**2%**
1.	ballot-level comparison	42	62	124	310
2.	ballot polling, 50% chance to complete*	150	330	1,310	8,200
2.	ballot polling, 75% chance to complete	260	590	2,330	14,500
2.	ballot polling, 90% chance to complete	440	980	3,850	24,100
2.	ballot polling, 95% chance to complete	580	1,310	5,100	32,300
3.	batch comparison	21,300 *17 batches*	31,000 *25 batches*	57,300 *48 batches*	125,700 *117 batches*

** sample size needed for an estimated 50% chance of attaining the risk limit in the first round*

Technical Notes for Table 1

The total number of voted ballots does not impact the number of ballots that need to be audited for Ballot Comparison or Ballot Polling Methods. For Batch Comparison, we used the number of ballots (464,144) and batches (3007) from the 2016 Presidential election.

Ballot-level Comparison: The estimates allowing for one one-vote overstatement (o1 = 1). See Lindeman and Stark, Gentle Introduction p. 4 (Equation 1).

Ballot Polling: These estimates are derived from simulation studies of the BRAVO method, assuming that every ballot bears a vote for one of the two candidates. (The BRAVO method is similar to the method described in Gentle Introduction, setting the tolerance t to 0, except for Steps 6 and 7. The criterion for stopping the audit, for a 10% risk limit, is $T > 10$ instead of $T > 9.9$, and there is no rule for whether and when to perform a full manual count.) Each estimate is based on 100,000 simulated audits. See Lindeman and Stark, Gentle Introduction pp. 3-4.

Ballot polling is less predictable than the comparison methods: even if the original count is highly accurate, an unlucky sample may not provide strong evidence for the reported outcome. A larger sample provides a greater chance of meeting the risk limit and complete the audit in just one round, without having to audit additional ballots. In Table 1, for instance, the “50 percent chance to complete” row shows sample sizes that give a 50 percent chance to complete the audit in one round. The table shows that to increase that chance to 95 percent generally requires auditing about four times as many ballots. Audit planners should take this tradeoff into account when setting the initial sample size.

Batch comparison: The number of batches to draw is based on section 3 in Stark (2009).[63] A convenient approximation for a 10% risk limit is $2 + 2.3 / m$, where m is the proportional margin. The estimate for number of votes audited comes from 2016 Rhode Island general election data. We treat each precinct’s in-person votes as a batch and each municipality’s absentee votes as a batch, and assume that the probability of selecting any batch is proportional to the number of cast votes it contains.

63. Philip B. Stark, “Efficient post-election audits of multiple contests: 2009 California tests,” refereed paper presented at the 2009 Conference on Empirical Legal Studies, preprint available at http://ssrn.com/abstract=1443314.

Risk-Limiting Audit Costs by Method and Contest Margin

Table 1 provides a summary comparison of the number of ballots that would need to be audited for different RLA methods and contest margins. The next series of three tables (below) estimate costs for major steps that each RLA method requires, and how those costs vary with differing contest margins.

In order to help election officials, policy makers, and the general public explore and understand costs associated with different types of risk-limiting audits, we used data from the Rhode Island pilot audits and other sources to create an interactive tool based on three related spreadsheets. **Tables 2A**, **2B** and **2C** contain illustrative snapshots of the three spreadsheets from the inter-active tool. We are providing an initial Excel implementation of the interactive tool to the Rhode Island Board of Elections. This version includes data and assumptions that are specific to current Rhode Island elections.

Table 2A contains user-settable parameters and set-up costs to prepare paper records for auditing that do not vary with different contest margins. At the top of **Table 2A**, users of the interactive tool can set three parameters: the average labor cost in dollars per hour, the total number of ballots expected to be cast for the election, and the average number of sheets per ballot. For this snapshot table, we have chosen:

- $20 per hour is the average labor cost per hour;
- 500,000 is the estimated total number of regular ballots that will be voted in precincts on election day in November 2020 (i.e., not counting mail ballots and provisionals that will be processed centrally in any case); and
- 1.5 is the estimated average number of sheets per ballot across all Rhode Island municipalities in November, 2020 (some jurisdictions within the state may require more than one sheet to accommodate all the contests and candidates for the 2020 election in that jurisdiction.[64]

The next set of rows in **Table 2A** show set-up costs that are common to all three types of risk-limiting audits (e.g., organizing shelving and labeling ballot containers).

The bottom three sets of rows in **Table 2A** show set-up costs that are specific to each RLA method – e.g., creating different types of ballot manifests for each method. For each activity shown in a particular row, there are columns whose cells contain the type of unit for that activity (e.g., ballot boxes or individual ballot sheets), the total number of such units, and the estimated number of person hours required to do that activity for each individual unit. Other columns to the right for person hours and cost are calculated from “# of units” and “pers hrs per unit”.

64. The numbers in these spreadsheets also assume that we are auditing 3 contests to a risk limit of 10%. In the future these numbers also could be incorporated as additional over-all parameters that could be set by users of the interactive tool.

Table 2A: Estimated Setup Costs for Three Different Methods Conducted Centrally

Labor Cost $ Per Hour =	**$20**	(This number applies throughout the tables and can be adjusted.)
Number of In-Person Ballots =	**500,000**	(The number of absentee ballots has little effect on audit costs in Rhode Island for 2020.)
Average Sheets Per Ballot =	**1.50**	(This affects scan time, sample size, and adjudication time.)
		[in the future, we might also allow users to specify risk limit and/or number of contests]

Steps & Sub-Steps

Steps in Common for All Three Audit Methods	Unit Type	# of Units	Pers hrs per unit	Person hours	Cost	Addl Est	Estimate Source & Notes
Organize shelving for 1,200 ballot boxes	box	1,200				$7,200	staff: $6/ballot box, 2 days for entire state
Transport ballot boxes to Providence	box	1,200		42	$840		staff: 3 people, 2 days, 7 hrs per day
Verify, sort, label and shelve all ballot boxes *(See steps by individual method below)*							
Generate random seeds and pull sheets				0.93	$19		RI pilot timing-4 people for 14 minutes
Subtotal (used in Total Costs Table)				**57**	**$8,339**		

1. Ballot -Level Comparison Method	Unit Type	# of Units	Pers hrs per unit	Person hours	Cost	Addl Est	Estimate Source & Notes
Rescan & create batches and manifest	sheet	750,000	0.0020	1,500	$30,000		Jan pilot* - may require diff labor pay rate?
Set up A/V and display				0			staff: TBD but not significant
Subtotal (used in Total Costs Table)				**1,500**	**$30,000**		

2. Ballot Polling Method	Unit Type	# of Units	Pers hrs per unit	Person hours	Cost	Addl Est	Estimate Source & Notes
Create manifest	ballot box	1,200	0.02	20	$400		use certificate ballot counts for each precinct
Set up A/V and display				0			staff: TBD but not significant
Subtotal (used in Total Costs Table)				**20**	**$400**		

3. Batch Comparison Method	Unit Type	# of Units	Pers hrs per unit	Person hours	Cost	Addl Est	Estimate Source & Notes
create batch manifest (per individual DS200)	DS200	559	0.01	6	$112		use certificate ballot counts for each precinct**
Subtotal (used in Total Costs Table)				**6**	**$112**		doesn't vary with margin

Table 2A Notes

* Row 17: From the pilot, 6 staff members can process about 3,000 ballots/hour = 500 ballots/person-hour or 0.002 person-hr/ballot (F17). It may be feasible to use fewer staff members.

** Row 24: Number of scanners from spreadsheet with DS200s per 2000 voters. Time per unit to create manifest is assumed same as for ballot polling.

Table 2B: Estimated Execution Costs by Method and Contest Margins (10% Risk Limit & 3 Contests)

				Different Contest Margins												
				10%			5%			2%			1%			
	Steps & Sub-steps	Unit	Pers Hr/ Unit	# of Units*	Pers Hr	Cost @Labor $/ Hr***	# of Units*	Pers Hr	Cost @Labor $/ Hr***	# of Units*	Pers Hr	Cost @Labor $/ Hr***	# of Units**	Pers Hr	Cost @Labor $/ Hr***	Estimate Source & Notes
	① Ballot-Level Comparison Risk-Limiting Audit Method															
	Handle boxes	box	0.20	93	19	$372	186	37	$744	465	93	$1,860	930	186	$3,720	E6: see note
	Pull sample	sheet	0.034	93	3	$63	186	6	$125	465	16	$313	930	31	$627	E7: see note
	Adjudicate	ballot	0.01	62	1	$14	124	1	$28	310	3	$69	620	7	$138	E8: see note
	Subtotal				**22**	**$448**		**45**	**$897**		**112**	**$2,242**		**224**	**$4,484**	
	② Ballot Polling Risk-Limiting Audit Method															
50% chance to complete	Handle boxes	box	0.20	990	198	$3,960	1,200	240	$4,800	1,200	240	$4,800	1,200	240	$4,800	see row 12 note
	Pull sample	sheet	0.035	495	17	$347	1,965	69	$1,376	12,300	431	$8,610	49,200	1,722	$34,440	see row 13 note
	Adjudicate	ballot	0.01	330	4	$73	1,310	14	$291	8,200	91	$1,820	32,800	364	$7,282	same times as for ballot comparison
	Subtotal				**219**	**$4,380**		**323**	**$6,466**		**762**	**$15,230**		**2,326**	**$46,522**	
75% chance to complete	Handle boxes	box	0.20	1,200	240	$4,800	1,200	240	$4,800	1,200	240	$4,800	1,200	240	$4,800	
	Pull sample	sheet	0.035	885	31	$620	3,495	122	$2,447	21,750	761	$15,225	87,150	3,050	$61,005	
	Adjudicate	ballot	0.01	590	7	$131	2,330	26	$517	14,500	161	$3,219	58,100	645	$12,898	
	Subtotal				**278**	**$5,550**		**388**	**$7,764**		**1,162**	**$23,244**		**3,935**	**$78,703**	
90% chance to complete	Handle boxes	box	0.20	1,200	240	$4,800	1,200	240	$4,800	1,200	240	$4,800	1,200	240	$4,800	
	pull sample	sheet	0.035	1,470	51	$1,029	5,775	202	$4,043	36,150	1,265	$25,305	144,000	5,040	$100,800	
	Adjudicate	ballot	0.01	980	11	$218	3,850	43	$855	24,100	268	$5,350	96,000	1,066	$21,312	
	Subtotal				**302**	**$6,047**		**485**	**$9,697**		**1,773**	**$35,455**		**6,346**	**$126,912**	
	③ Batch Comparison Risk-Limiting Audit Method (for 3 contests)															
	Handle boxes	batch	0.24	25	6	$120	48	12	$230	117	28	$562	232	56	$1,114	see row 27 note
	Batch sheets	sheet	0.003	31,000	102	$2,046	57,300	189	$3,782	125,700	415	$8,296	211,600	698	$13,966	see row 28 note
	Count/record	ballot	0.013	31,000	388	$7,750	57,300	716	$14,325	125,700	1,571	$31,425	211,600	2,645	$52,900	see row 29 note
	Subtotal				**496**	**$9,916**		**917**	**$18,337**		**2,014**	**$40,283**		**3,399**	**$67,979**	

Table 2B Notes

* For each margin, the number of units required is from Exhibit 1 of the RI report.

** For margins of 1% Mark Lindeman provided the required number of units.

*** The labor costs are entered as a parameter on the Setup Cost Spreadsheet.

Table 2B Notes (cont.)
E6: For retrieving and replacing boxes, Miguel Nunez estimates using a small pallet with ten boxes, 20 minutes to retrieve per pallet, 10 mins to return, 2 mins to retrieve per box and 1 minute to return (20 + 10 + 2*10 + 1*10) = 60 minutes per pallet or 6 minutes (0.1 hours) per box to retrieve and return. **2 staff, so 0.2 person hrs/box** (not physically necessary). Box count for 1% is conservative.
E7: Jan pilot 61 seconds median to pull sheet from a new box. 2 persons working so 2.02 minutes (0.0337 hours) per sheet
E8: For three contests, based on Jan pilot avg time to evaluate one contest: 25 seconds, ten contests, 62 seconds, three contests: 40 seconds (.0111 hours). (It takes a few seconds to handle ballots on which no audited contests appear, but the # of ballots is more relevant than the # of sheets.)
Row 12: Assume 2 boxes per retrieved sheet with maximum 1,200.
Row 13: From pilot, using scale method: 1.05 minutes for all new box. For 2 persons working, 1.05 * 2 / 60 = 0.035 person hours per ballot.
Row 27: Each batch will have multiple boxes stored near each other. Assume that collecting all the boxes for a batch takes 1.2 times the amount of time to collect one box (set in row 6).
Row 28: Use of sheet is conservative: depending on contests, additional sheets may be set aside without batching. Pilot timings include start-up issues; the timing here is the fastest observed, 0.10 min * 2 persons/60 = 0.003 person hours per unit.
Row 29: Pilot. .17 min was medium observed amount for evaluation and tallying; with 4 persons, 0.17 * 4 / 60

The spreadsheet table in **Table 2B** shows the costs of retrieving and evaluating sampled ballots. Once again, rows show different types of activities for each method under the heading for that method, while columns show type of unit, number of units, and person-hours per unit.

At the right side of **Table 2B**, there are four sets of three cells each: number of units, calculated person hours and cost; each set corresponds to a different contest margin (10%, 5%, 2%, 1%).

Table 2B shows that the closer the margin for an audited contest, the more units (boxes and ballot sheets) need to be handled and counted. The number of units to be audited is determined primarily by the contest margin, rather than the total number of votes cast for a particular contest being audited.

In **Table 2B**, note that the ballot-polling RLA method activities (handle boxes, pull sample, and adjudicate) are grouped into three vertical sets of different “chances to complete” – 50%, 75%, and 90%. For example, **Table 2B** shows that to attain a 10% risk limit using the ballot-polling method for a contest where the winning margin was only 1%, there would be a 75% chance of completing the audit in the first round by examining 87,150 ballot sheets from 58,100 ballots.

Unlike ballot-level comparison and batch comparison methods, the sample size needed to attain a specified risk limit in the first round of auditing (assuming only minor discrepancies) is quite variable for a ballot polling audit. For instance, if by chance the randomly chosen sample includes many more ballots with the loser selected than appear on average in all the ballots, many more ballots would require examination than would be required on average. So instead of giving a single “# of units” (as shown for the other two RLA methods), the spreadsheet shows three different sets of “# of units” that need to be in the sample in order to achieve a certain percentage probability of completing the audit with that size sample (and assuming that no discrepancies are found).

Table 2C: Combined Total Costs

		Different Contest Margins								
		10%		5%		2%		1%		
	RLA Methods	Person Hours	Cost	Person Hours	Cost	Person Hours	Cost	Person Hours	Cost	
	① Ballot -Level Comparison Method									
	Set up in common	57	$8,339	57	$8,339	57	$8,339	57	$8,339	From Setup Table
	Set up for method	1,500	$30,000	1,500	$30,000	1,500	$30,000	1,500	$30,000	
	Execution	22	$448	45	$897	112	$2,242	224	$4,484	From Execution Table
	Total	**1,579**	**$38,787**	**1,602**	**$39,235**	**1,669**	**$40,581**	**1,781**	**$42,823**	
	② Ballot Polling Method									
50% chance to complete	Set up in common	57	$8,339	57	$8,339	57	$8,339	57	$8,339	From Setup Table
	Set up for method	20	$400	20	$400	29	$400	20	$400	
	Execution	219	$4,380	323	$6,466	762	$15,230	2,326	$46,522	From Execution Table
	Total	**296**	**$13,118**	**400**	**$15,205**	**838**	**$23,969**	**2,403**	**55,260**	
75% chance to complete	Set up in common	57	$8,339	57	$8,339	57	$8,339	57	$8,339	From Setup Table
	Set up for method	20	$400	20	$400	20	$400	20	$400	
	Execution	278	$5,550	388	$7,764	1,162	$23,244	3,935	$78,703	From Execution Table
	Total	**354**	**$14,289**	**465**	**$16,502**	**1,239**	**$31,983**	**4,012**	**$87,442**	
90% chance to complete	Set up in common	57	$8,339	57	$8,339	57	$8,339	57	$8,339	From Setup Table
	Set up for method	20	$400	20	$400	20	$400	20	$400	
	Execution	302	$6,047	485	$9,697	1,773	$35,455	6,346	$126,912	From Execution Table
	Total	**379**	**$14,785**	**562**	**$18,436**	**1,850**	**$44,194**	**6,423**	**$135,651**	
	③ Batch Comparison Method									
	Set up in common	57	$8,339	57	$8,339	57	$8,339	57	$8,339	From Setup Table
	Set up for method	6	$112	6	$112	6	$112	6	$112	
	Execution	496	$9,916	917	$18,337	2,014	$40,283	3,399	$67,979	From Execution Table
	Total	**558**	**$18,366**	**979**	**$26,788**	**2,077**	**$48,733**	**3,461**	**$76,430**	

The spreadsheet/table in **Table 2C** brings together sub-totals from the setup and execution tables (**Tables 2A** and **2B**) and summarizes for each method the common set-up costs, method-specific set-up costs, and execution costs, along with totals for each. The sub-totals and totals are grouped and broken down by method as well as for different contest margins, as in **Table 2B**.

One of the main take-away messages from the summary table in **Table 2C** is that total costs for ballot-comparison audits increase only slightly with smaller contest margins, while the costs for the other two methods increase dramatically as margins decrease. Moreover, it can become very expensive to complete an audit using the ballot-polling method when contest margins are small.

Summary of Policy Choices for the Rhode Island Board of Elections[65]

When considering what recommendations to make to Rhode Island, the working group asked itself five questions:

1. *Which of the three methods – ballot-level comparison, ballot polling, or batch comparison – should Rhode Island adopt?*
2. *If ballot polling, which retrieval method – count, scale, ruler or k-cut – should the state use?*
3. *Should the audits be conducted centrally or should they be distributed?*
4. *What software should be used?*
5. *What timelines would make Rhode Island best situated to conduct its first official risk-limiting audit in April 2020?*

Recommendations

Based on our experience and data from Rhode Island's very informative January pilot risk-limiting audits, analysis of data from those pilot audits and other sources, and discussions with Rhode Island Board of Elections staff, the members of the RIRLA Working Group offer the following recommendations:

1. Implement a ballot-level comparison risk-limiting audit.

While there are a variety of reasons to support a ballot-level comparison risk-limiting audit, four are worth highlighting:

1. Ballot-level comparison audits enable officials to trace discrepancies to the individual ballot. In the event that something goes wrong, this audit method can provide more information about the performance of the voting system than either the ballot polling or batch comparison approach. Such information is key for the continuous improvement of the election and auditing processes.

2. Ballot-level comparison audits are particularly attractive in Rhode Island because the work required for performing any recount could also be used for the audit. By Rhode Island law, candidates have a right to a recount in close margin contests,[66] and the recount simply consists of refeeding all of the ballots into the optical scanners. So far, the state has yet to conduct a refeed for a statewide contest. However, by adopting ballot-level comparison audits, which requires a rescan of all ballots, Rhode Island would be prepared for the eventual statewide recount scenario.

65. The consensus recommendations in this section reflect the opinions of the following members of the Rhode Island RLA Working Group (in alphabetical order): Wilfred Codrington, Lynn Garland, Mark Lindeman, John Marion, John McCarthy, Ron Rivest, and Luther Weeks. Board of Elections staff members Miguel Nunez and Steve Taylor participated actively in Rhode Island RLA Working Group discussions but they did not feel it would be appropriate for them to join in making recommendations because the Board of Elections staff may be asked to make their own independent recommendations.
66. http://webserver.rilin.state.ri.us/Statutes/TITLE17/17-19/17-19-37.1.HTM. In races where there are fewer than 20,000 votes a recount occurs when the margin is less than 2% or 200 votes, whichever is less. In races between 20,001 and 100,000 votes a recount occurs when the margin is less than 1% or 500 votes, whichever is less. In races where there are 100,001 or more votes a recount occurs when the margin is less than 0.5% or 1,500 votes, whichever is less.

3. Ballot-level comparison audits enable opportunistic audits to easily gain some confidence in the outcomes of all the contests.This method is easiest to extend to both statewide and non-statewide contests. As the pilot demonstrated, conducting a ballot-level comparison risk-limiting audit of a large number of statewide contests (ten in the pilot) need not take much longer than auditing just one contest, depending on the margins in the contests. Risk-limiting audits of smaller contests can readily be added as well. Moreover, additional contests can be opportunistically audited on some or all of the ballots in the audit sample, thus easily gaining additional evidence about the accuracy of the vote counts. Extending the other audit methods to multiple contests typically is far more laborious and/or less effective.

4. The ballot-level comparison method has the most predictable workload and cost. To be clear, because of the rescanning requirement, the ballot-level comparison method would have relatively large upfront costs (at least using the state's current voting systems). Those costs are fairly predictable, however, because they are largely determined by the number of hours required to conduct the rescan, which one can conservatively estimate before the election based on expected voter turnout. (Costs are minimal after the rescan unless the audit escalates). The workload for the ballot-level comparison audit is dependent on voter turnout, a relatively stable figure. However, the workload for the alternatives is unpredictable because it depends much more on the contest margin, which varies by contest and election year. For instance, as Table 1 shows, the difference in workload between a 10-point margin and a 2-point margin (at a 10% risk limit) is almost 100,000 hand-counted ballots if batch comparison is used, but only about 250 ballots if ballot-level comparison is used.

We believe it is feasible to implement statewide ballot-level comparison in the 2020 election cycle. That said, ballot-level comparison is not an all-or-nothing proposition. For the presidential preference primary or the general election, the Board of Elections might choose to retabulate part of the state's ballots, and audit these ballots using ballot-level comparison, using ballot polling or batch comparison for the remaining ballots. (The results from these two "strata" could be combined to produce the overall audit result.) With careful planning, the audit can be designed so that if the retabulation goes faster or slower than expected, the audit can adapt accordingly. The Board then can adjust its audit plan for November 2020. Even if the Board decides that it is premature to retabulate all ballots after the presidential primary or general election, it can use retabulation and ballot-level comparison to implement risk-limiting audits in selected non-statewide contests, while using one of the other methods for the presidential and other statewide contests.

If the Board decides not to use retabulation and ballot-level comparison statewide, the Working Group has no firm consensus on which alternative method is preferable. However, we caution that as shown in table 1, ballot polling can become unwieldy when a target contest has a small margin.

Establish objective criteria for which races will be audited.

In April 2020 the Board will have to audit the Presidential Preference Primary contests. The Board has significant discretion to determine which contests to as part of the November 2020 election. The Working Group recommends the Board audit the presidential contest and establish by regulation objective criteria and a transparent process for determining additional contests to be audited.

Conduct a centralized audit.

Although there are some advantages to having local election officials conduct the audit in a decentralized manner (and it may be a good approach in future years), Rhode Island's first official risk-limiting audits should be conducted centrally. Centralization facilitates the training, communications, and management required for a successful audit.

Both centralized and decentralized audits have benefits from the observability standpoint. However, conducting ballot-level comparison audits centrally has the added advantage of permitting election integrity advocates and the public to watch the entire process. Spreading the audits among more locations (up to all 39 of Rhode Island's cities and towns would make it less likely that trained observers would always be present at each location.

Consult local election officials.

Before making any final decisions, particularly ones that may have a great impact on local election officials (e.g. requiring them to carry out a decentralized risk-limiting audit), the Board of Elections should consult those local officials and seek their input. Local election officials may perceive the additional work as an unfunded mandate. (Notably, a centralized audit may be the best approach for enlisting participation, as local election officials who are willing and able could opt to participate in the RLA – much as Michigan successfully recruited municipal election officials to participate in its countywide RLA pilots of June 2019.)

Conduct a practice audit.

The Board of Elections should conduct another (semi-private) pilot audit in the fall of 2019. This is particularly important given the likelihood that Rhode Island will use a different (or at least substantially revised) software tool. A second pilot would be even more warranted should the Board choose to employ methods not tested during the January 2019 pilots (such as using a hybrid method or conducting an audit in a decentralized manner).

Use Arlo audit software.

Although the Audit Conductor software was adequate for the January pilot, the Board of Elections should use the Arlo software from Voting Works.[67] This open-source software was originally developed to support Colorado's statewide risk-limiting audits. The Arlo software supports parallel input of data from multiple audit adjudicator boards simultaneously in centralized or distributed audit locations. By becoming a paying customer, Rhode Island can help determine future software enhancements (including those scheduled for the coming months) and would get documentation and assistance from paid professionals (who have provided support to Colorado state and local election officials.

67. https://voting.works/rla/.

Appoint an ongoing expert advisory council.

Beginning in 2020 the Rhode Island Board of Elections will have to conduct the audits themselves. We believe that this will require the assistance of experts from inside and outside of the state. The Board should appoint a council of experts to advise them in this process. It might look to the example of Colorado, which has engaged a group of local election officials, auditing experts, and others, to help develop and review audit protocols, rules and regulations, and software, among other things.

Initiate rulemaking.

The Board of Elections must initiate rulemaking as soon as practicable. Rulemaking is necessary for the Board to make key policy choices, which include:

- Establishing ballot interpretation rules;
- Determining contests subject to audit;
- Setting risk limits;
- Adjusting the election calendar
- Harmonizing audits processes with recounts/refeeds requirements Clarifying public notice requirements and rules governing access and observation (as required by Rhode Island's RLA statute); and
- Specifying qualifications for serving as an election judge.

Develop schedule with milestones.

The Rhode Island Board of Elections needs to develop and publish the major milestones they must meet to implement risk-limiting audits. Those milestones include choosing an audit method, conducting additional pilots, and completing the regulatory process, and should account for the period between July 2019 through the general election in November 2020.

Endorse vendor recommendations.

The Rhode Island Risk-limiting Audits Working Group plans to make recommendations to the election machine vendor (as summarized in Appendix A), and urges the Rhode Island Board of Elections to send its own letter to the vendor making similar recommendations.

✓ Part V:
Conclusion

The Rhode Island risk-limiting audit pilots exceeded expectations. They provided an opportunity for election administrators to gain hands-on experience with RLAs. They gave theoreticians an opportunity to test new methods. They gave advocates another data point to point to when arguing that RLAs are possible for any voting system.

As the first known effort to test all three types of RLAs simultaneously, and to provide comprehensive timing and cost measurements, we believe these pilots provide valuable insights for administrators, theoreticians and advocates alike. While we endeavored to test many different methods, there is much more to learn and we look forward to future pilots in Rhode Island and elsewhere.

We hope this report will be a useful tool for the Rhode Island Board of Elections as it drives toward implementing RLAs beginning in April 2020. We thank their staff who have worked with us over the past seven months to plan, implement, measure and now describe the pilot risk-limiting audits.

Glossary

Audit judges
Officials participating in performing the audit.

Ballot manifest
A catalog prepared by election officials listing all the physical paper ballots and their locations in sequence.

Ballot sheets
A single piece of paper that forms part of a paper ballot. Paper ballots may contain multiple sheets.

Batch ID
A unique ID associated with a batch of ballots, used for labeling the batch and for identifying the batch in the ballot manifest.

Common Data Format (CDF)
A format specification designed to allow a kind of data, such as Cast Vote Records, to be interoperably transferred between systems (for instance, from election management systems to audit software).

Compliance audits
Audits that evaluate the compliance with laws, procedures, or standards. E.g. evaluating the security of ballot storage or the conduct of officials on election day.

Counting scale
A scale that can estimate counts of objects (such as ballots) based on the objects' piece weight.

Direct Recording Electronic (DRE)
A vote-capture device that allows electronic presentation of a ballot, electronic selection of valid contest options, and electronic storage of contest selections as individual records. It also provides a summary of these contest selections.

Election certificates
Documents submitted by pollworkers that report the counts of cast and voided ballots, and of voters, in each precinct.

Hashmarking
A method of counting votes where for each vote counted one person reads the vote, and one or more people make tally marks is sets of five marks for each candidate or yes/no vote.

Help America Vote Act (HAVA)
A U.S. Congressional Act in 2002 which provided funds and guidance to states in acquiring new voting equipment.

Jogger
A mechanical device that vibrates vigorously to reduce misalignment among ballot sheets or other pieces of paper.

Logic and accuracy testing
Equipment and system readiness tests whose purpose is to detect malfunctioning devices and improper election-specific setup before the equipment or systems are used in an election. Election officials conduct L&A tests prior to the start of an election as part of the process of setting up the system and the devices for an election according to jurisdiction practices and conforming to any state laws.

MOVE Act
The Military and Overseas Voter Empowerment Act of 2009, federal legislation to facilitate voting by military and overseas voters.

Opportunistic auditing
Auditing of additional contests on selected ballots beyond the contests for which risk limits have been specified.

Optical scan
Voting system that counts votes marked in contest option positions on the surface of a paper ballot.

Outstack
During ballot scanning, to direct some ballots into a separate bin, apart from the main stack of scanned ballots.

Overvote
Occurs when the number of selections made by a voter in a contest is more than the maximum number allowed.

Precinct
Election administration division corresponding to a geographic area that is the basis for determining which contests the voters legally residing in that area are eligible to vote on.

Presidential preference primary
Primary election in which voters choose the delegates to the presidential nominating conventions allotted to their states by the national party committees.

Provisional ballot
A failsafe ballot provided to a voter whose eligibility for a regular ballot cannot be immediately determined. The ballot may be counted or further processed depending on state law.

Pseudorandom number generator
Software algorithm that, given an initial seed value, generates a sequence of numbers that approximate the properties of random numbers (for instance, the next pseudorandom number cannot be predicted from any previous numbers in the sequence) but can be reproduced if one knows the seed. If the initial seed is a random seed, the sequence can be used as a random sample.

Random seed
A randomly generated number that provides the initial seed (input) to a pseudorandom number generator.

Risk-limiting audits (RLAs)
Procedure for checking a sample of ballots (or voter verifiable records) that is guaranteed to have a large, pre-specified chance of correcting the reported outcome if the reported outcome is wrong (that is, if a full hand count would reveal an outcome different from the reported outcome).

Serial ID number
A number imprinted on each ballot sheet during scanning that allows the sheet to be identified and associated with the corresponding Cast Vote Record. The DS850 used in the pilot imprinted nine-digit IDs, the first of which was 237000001.

Simple random samples
Random samples in which every item (such as a ballot sheet) has an equal and independent probability of being selected.

Sort-and-stack method
A method of counting votes where ballots are sorted into stacks by selections in contests and then the number of ballots in a stack are counted to determine the number of votes for a candidate or yes/no.

Transitive RLA
An RLA in which the ballots are retabulated, obtaining a Cast Vote Record for each ballot sheet, and the audit uses these new Cast Vote Records instead of data from the original tabulation. If the retabulation reports the same outcome(s) as the original tabulation, an RLA of the retabulation can confirm the original outcome(s).

Voting system

Equipment (including hardware, firmware, and software), materials, and documentation used to define elections and ballot styles, configure voting equipment, identify and validate voting equipment configurations, perform logic and accuracy tests, activate ballots, capture votes, count votes, reconcile ballots needing special treatment, generate reports, transmit election data, archive election data, and audit elections.

Write-in

A type of contest option that allows a voter to specify a candidate, usually not already listed as a contest option. Depending on election jurisdiction rules, in some cases only previously approved names will be considered as valid write-in contest selections.

Appendices

Appendix A: Recommendations to the Vendor

Develop a mechanism to imprint a unique, pseudorandom number on each ballot immediately upon scanning and to store this number in the corresponding cast vote record (CVR).

Ballot-level comparison RLAs typically involve examining the fewest number of ballots, but they require individual cast ballots to be linked with a one-to-one association to each individual corresponding cast vote record (CVR). This linkage can be achieved by imprinting a unique pseudorandom number on the physical ballot and including this number in that ballot's CVR. In Rhode Island, most voters cast their ballots in-person using a DS200 scanner which currently lacks this capability. Therefore, conducting a ballot-level comparison audit requires officials to re-scan the ballots and imprint each one during the second scan. (For the pilot, we used the DS850 central scanner for the second scan, which has imprinting capability. In this case, the numbers did not have to be random because the ballots were no longer associated with individual voters.) Rescanning and imprinting adds time and cost to the audit, and the additional step creates more room for mishandling and error.

The vendor should add this functionality to in-precinct voting equipment, so that unique, pseudorandom numbers will be generated for and imprinted on each ballot after the scan. It is important that the pseudorandom number is added only after the ballot is scanned and accepted by the machine. (This prevents a number from being imprinted on a ballot that is invalid and returned to a voter only to be re-imprinted with a different number.) It is likewise essential that the pseudorandom numbers be generated in a manner that ensures voter anonymity; they must be completely dissociated from individual voters, and ensure that no one can match the voter's identity to the voter's cast ballot.

Mike Goetz, Vice President of Product Management at ES&S, said that the company explored the possibility of retrofitting the current model of the DS200 to add this functionality, but that those machines do not have sufficient physical space to add imprinting hardware. However, he said that the company was looking into adding this capability to the next generation DS200 model, which is currently under development.

Simplify and standardize the process for election officials to export the necessary data to support audits.

The pilot audits required considerable vendor support to export the voting data and to ensure its usability. While the Board and the Rhode Island RLA Working Group appreciated the vendor's willingness to provide assistance, such information to support audits should be readily available for exporting CVRs and other information from the Election Management System (EMS) without special assistance, especially as more and more jurisdictions conduct audits as standard practice.

Rhode Island election officials (and any agents or consultants who assist with audit design, planning, and implementation) must be able to retrieve essential data in a usable format more readily. We recommend that the vendor adopt the NIST-VVSG interoperable Common Data Format (CDF) standard for CVRs to provide a uniform export for audits. (See Appendix O for the format used in Rhode Island.)

Continue to support Automark ballot-marking device systems.

The Automark system created ballots that were easy to incorporate in the Rhode Island pilot audit. Because of their similar size and appearance to hand-marked ballots, the Automark ballots presented little risk of jeopardizing ballot secrecy. (Notably, ballot security is both a general requirement for election administration and an essential component of ballot protection, a best practice for risk-limiting audits.) To ensure the best results for Rhode Island and the many other states that employ the Automark system, ES&S should continue to support the Automark even as the company makes system upgrades and other improvements.

Provide election officials with voting machines' criteria for counting a mark within an oval as a vote.

The vendor should make any voting machine criteria (including the percentage of oval that needs to be marked to be counted) available to election officials so they can be considered in analyzing discrepancies between Cast Vote Records and audit interpretations. The vendor should also provide any available tuning parameters that can be adjusted for recognizing valid votes.

Make improvements to the sensitivity of the DS850.

During retabulation the DS850 repeatedly outstacked a significant number of ballots requiring them to be re-fed. The DS850 also jammed fairly frequently. We have heard that other jurisdictions that use DS850's have experienced similar problems. Further improvements to the DS850 that will allow them to outstack fewer ballots and have fewer jams will increase the efficiency of transitive audits. Until improvements are made ES&S should provide additional training and staffing to jurisdictions using DS850s for risk-limiting audits for free or at a discounted rate.

Appendix B: State Audit Working Group

The State Audit Working Group (SAWG) is an informal group of election integrity advocates, statisticians, computer scientists, election officials, and citizens who discuss and promote different kinds of election audits to increase election security and ensure the correctness of election outcomes. The SAWG was founded in early 2008 by Mark Halvorson from Citizens for Election Integrity Minnesota, Luther Weeks from Connecticut Voters Count, and John McCarthy and Pam Smith from Verified Voting, following the first Post Election Audit Summit meeting in October, 2007 in Minneapolis, Minnesota.

The SAWG has held a weekly meeting (sometimes bi-weekly) via teleconference, since early 2008. Weekly calls typically have anywhere from six to fifteen participants, and a larger number of participants communicate via an active email list. that currently includes some 60 individuals. Luther Weeks from Connecticut Voters Count took over convening the weekly calls from Mark Halvorson in 2013. SAWG has offered post election audit advice to states including Colorado, which was the first state to implement a statewide RLA.

Appendix C: H 5704 Sub A

2017 -- H 5704 SUBSTITUTE A

========
LC001568/SUB A
========

S T A T E O F R H O D E I S L A N D

IN GENERAL ASSEMBLY

JANUARY SESSION, A.D. 2017

A N A C T

RELATING TO ELECTIONS --POST-ELECTION AUDITS

Introduced By: Representatives Ajello, Regunberg, Chippendale, Craven, and Tanzi

Date Introduced: March 01, 2017

Referred To: House Judiciary

It is enacted by the General Assembly as follows:

SECTION 1. Chapter 17-19 of the General Laws entitled "Conduct of Election and Voting Equipment, and Supplies" is hereby amended by adding thereto the following section:

17-19-37.4. Post-election audits.

(a) The general assembly hereby finds, determines, and declares that auditing of election results is necessary to ensure effective election administration and public confidence in the election results. Further, risk-limiting audits provide a more effective manner of conducting audits than traditional audit methods in that risk-limiting audit methods typically require only limited resources for election contests with wide margins of victory while investing greater resources in close contests.

(b) Commencing in 2018 the board in conjunction with local boards is authorized to conduct risk-limiting audits after all statewide primary, general and special elections in accordance with the requirements of this section. Commencing in 2020 the state board in conjunction with local boards must conduct risk-limiting audits after the presidential preference primary, and general elections in accordance with requirements in this section.

(c) The audit program shall be conducted as follows:

(1) The state board shall determine what local, statewide and federal contests are subject to a risk-limiting audit;

(2) The state board shall provide notice pursuant of chapter 46 of title 42 of the time and place of the random selection of the audit units to be manually tallied and of the times and places

of the audits;

(3) The state board shall make available to the public a report of the vote tabulating device results for the contest, including the results for each audit unit in the contest, prior to the random selection of audit units to be manually tallied and prior to the commencement of the audit;

(4) The state board in conjunction with the local boards shall conduct the audit upon tabulation of the unofficial final results as provided in §§17-19-36 and 17-19-37; and

(5) The state board in conjunction with the local boards shall conduct the audit in public view by manually interpreting the ballots according to rules established by the state board in accordance with chapter 45 of title 42.

(d) If a risk-limiting audit of a contest leads to a full manual tally of the ballots cast using the voting system, the vote counts according to that manual tally shall replace the vote counts reported pursuant to §§17-19-36 and 17-19-3 7 for the purpose of determining the official contest results pursuant to §§17-22-5.2 and 17-22-6.

(e) For purposes of this section, the following terms have the following meanings:

(1) "Audit unit" means a precinct, a set of ballots, or a single ballot. A precinct, a set of ballots, or a single ballot may be used as an audit unit for purposes of this section only if all of the following conditions are satisfied:

(i) The relevant vote tabulating device is able to produce a report of the votes cast in the precinct, set of ballots, or single ballot.

(ii) Each ballot is assigned to not more than one audit unit.

(2) "Contest" means an election for an office or for a measure.

(3) "Risk-limiting audit" means a manual tally employing a statistical method that ensures a large, predetermined minimum chance of requiring a full manual tally whenever a full manual tally would show an electoral outcome that differs from the outcome reported by the vote tabulating system for the audited contest. A risk-limiting audit shall begin with a hand tally of the votes in one or more audit units and shall continue to hand tally votes in additional audit units until there is strong statistical evidence that the electoral outcome is correct. In the event that counting additional audit units does not provide strong statistical evidence that the electoral outcome is correct, the audit shall continue until there has been a full manual tally to determine the correct electoral outcome of the audited contest.

(4) "Unofficial final results" means election results tabulated pursuant §§17-19-36 and 17-19-37.

(f) The results of any audits conducted under this section shall be published on the website of the state board within forty-eight (48) hours of being accepted by the state board. If the audit involved a manual tally of one or more entire precincts, then the names and numbers of all precincts audited and a comparison of the vote tabulator results with the hand counts for each precinct shall be published with the audit results on the website.

(g) Any audit required under this section shall not commence for any election subject to a recount pursuant to §§17-19-37.1, 17-19-37.2, and 17-19-37.3 until the conclusion of said recount.

(h) The state board shall promulgate rules, regulations, and procedures in accordance with chapter 45 of title 42 necessary to implement this section.

SECTION 2. This act shall take effect upon passage.

========
LC001568/SUB A
========

EXPLANATION

BY THE LEGISLATIVE COUNCIL

OF

A N A C T

RELATING TO ELECTIONS --POST-ELECTION AUDITS

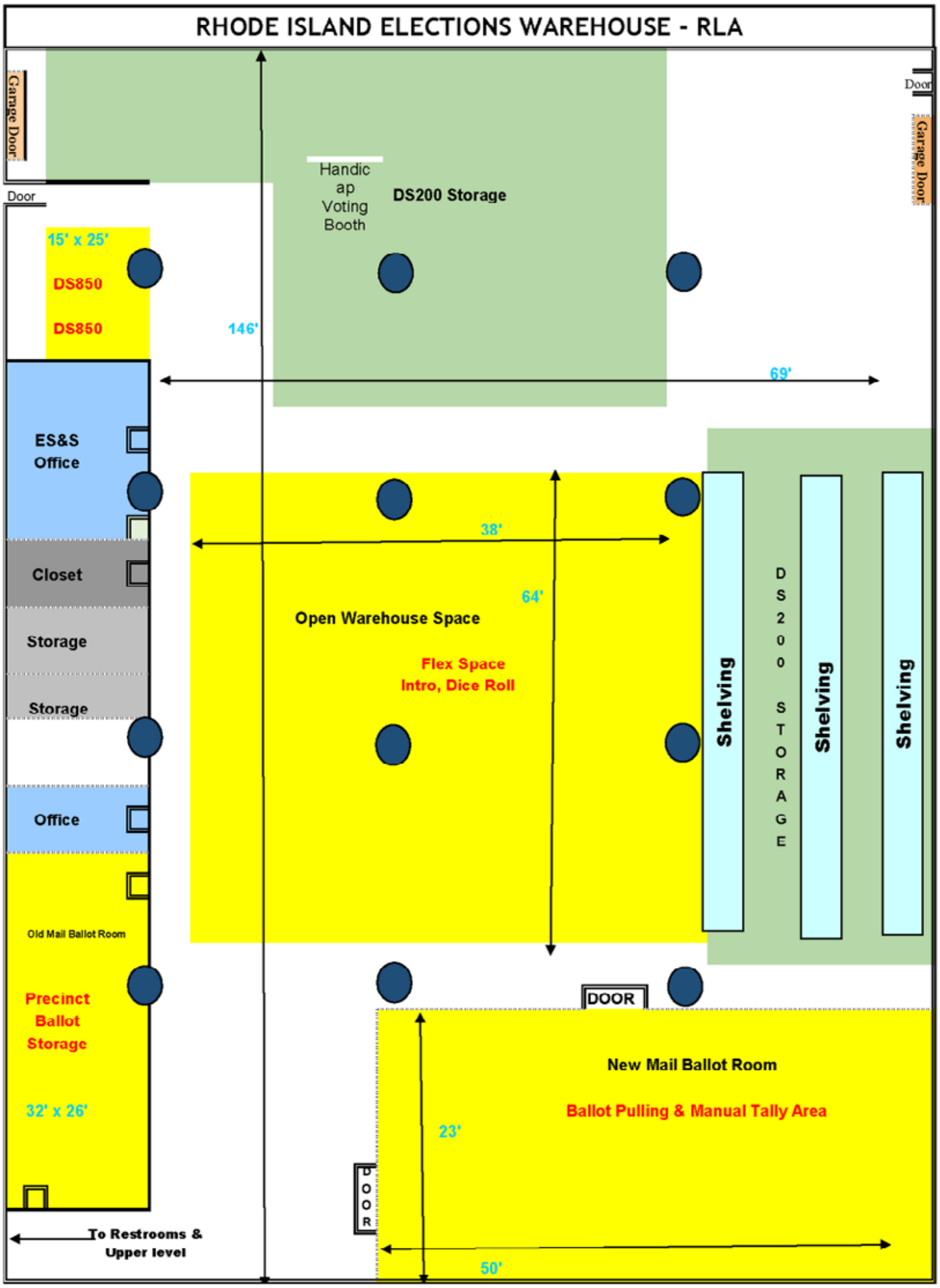
***

This act would authorize the board of elections to establish a post-election risk-limiting audit program to improve the accuracy of election results.

This act would take effect upon passage.

========
LC001568/SUB A
========

Appendix D: Pilot Audit Floor Plan



Appendix E:
Detailed Protocol Describing Each Phase of Pilot

Available at:
https://drive.google.com/file/d/1S65PJqht7EhdnqHujkx4IEjX1e_7u7M6/view?usp=sharing

Appendix F:
Ballot Hashmarking Sheet

Stack Number: ☐ **Ballots In Stack:** (count) ☐ **Batch:** ☐

☐ Write-ins ☐ Recount Number: ☐

		Up to 5 hashmarks per cell, e.g. 𝍸										Total
		5	10	15	20	25	30	35	40	45	50	Counts
Senator in Congress											**Total For Contest:**	
Whitehouse	DEM											
Flanders	REP											
NO VOTE												
Represenative in Congress											**Total For Contest:**	
Cicilline	DEM											
Donovan	REP											
NO VOTE												
Governor											**Total For Contest:**	
Raimondo	DEM											
Gilbert	MOD											
Fung	REP											
Armstrong	COM											
Munoz	IND											
Trillo	IND											
NO VOTE												

Appendix G:
Example of Ballot Manifest

	A	B	C	D	E	F	G	H
1	Municipality	Precinct Number	Box Letter	Folder Number	Batch ID	# of Sheets	First Imprinted ID	Last Imprinted ID
2	Bristol	201	A	1	0201-A-1	297	237000001	237000297
3	Bristol	201	A	2	0201-A-2	100	237000298	237000397
4	Bristol	201	A	3	0201-A-3	99	237000398	237000496
5	Bristol	201	A	4	0201-A-4	99	237000497	237000595
6	Bristol	201	A	5	0201-A-5	99	237000596	237000695
7	Bristol	201	A	6	0201-A-6	99	237000696	237000794
8	Bristol	201	A	7	0201-A-7	54	237000795	237000848
9	Bristol	202	B	1	0202-B-1	198	237000849	237001047
10	Bristol	202	B	2	0202-B-2	199	237001048	237001247

Appendix H:
Tracking Sheet

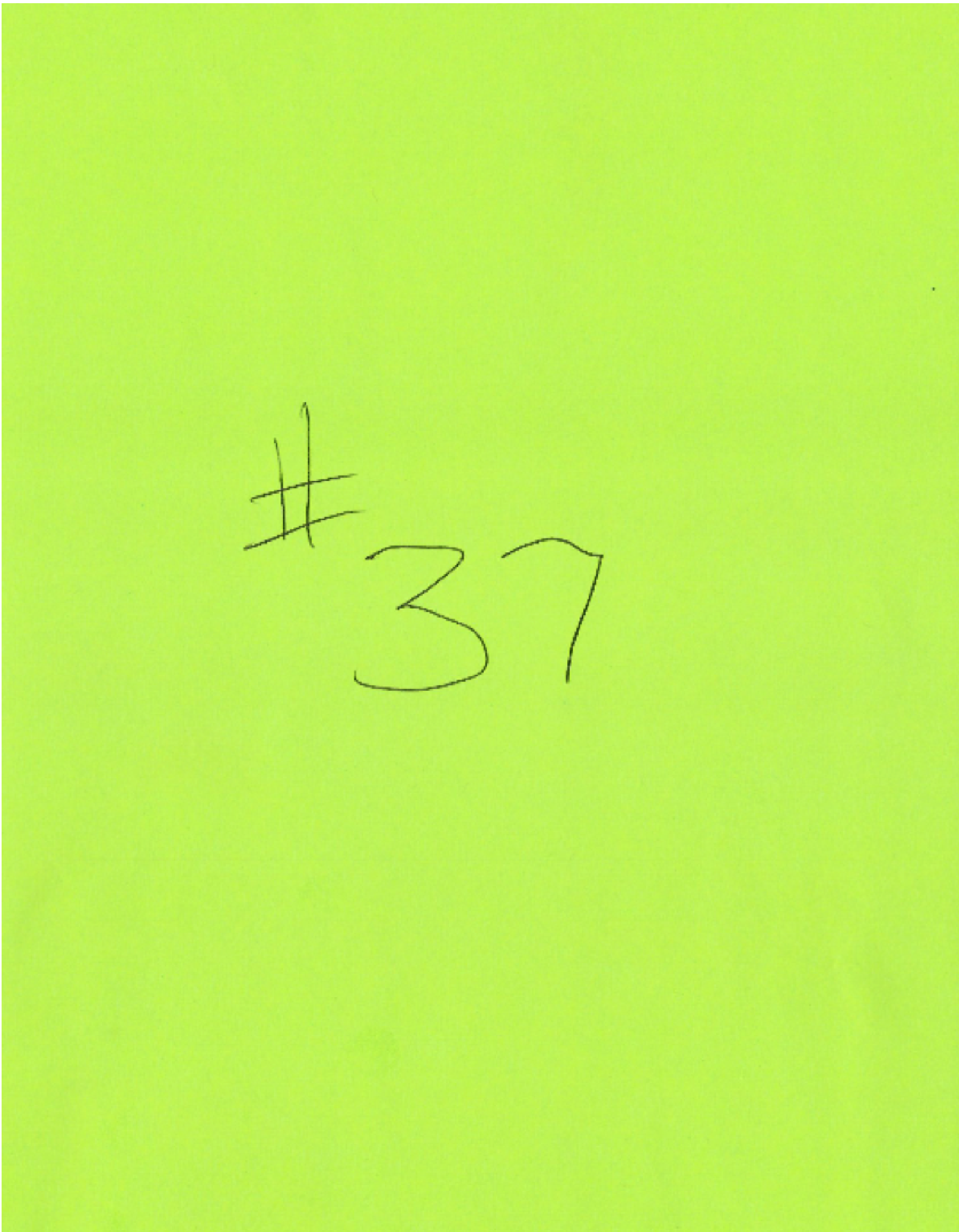
Precinct ID	
Box Letter	
Folder Number	
Number of Ballots	
Beginning ID Number	
Ending ID Number	

Precinct ID	
Box Letter	
Folder Number	
Number of Ballots	
Beginning ID Number	
Ending ID Number	

Prepared by __________ ___________

Prepared by __________ ___________

Appendix I: Placeholder Sheet



Appendix J: Cover Sheet

Logic ID

Precinct ID	
Box Letter	
Folder #	
Ballot #	

____________ ____________

____________ ____________

____________ ____________

Logic ID

Precinct ID	
Box Letter	
Folder #	
Ballot #	

____________ ____________

____________ ____________

____________ ____________

Appendix K: Pull Sheet

Portsmouth team 1 [2703]

Counting sample

batch 2703-C-3, #143
batch 2703-C-4, #207

Ruler sample

2703-C-1, #196
*** #7 from end ***
2703-C-2, #120 (19 mm)
2703-C-3, # 9 (count)
--AND-- #127 (20 mm)
--AND-- #168 (26 mm)
2703-C-4, # 55 (9 mm)
--AND-- #163 (26 mm)
2703-D-1, # 34 (5 mm)
--AND-- # 84 (13 mm)
--AND-- #111 (17 mm)

k-cut sample

batch 2703-C-1, # 42
batch 2703-C-2, # 44
--AND-- #128
--AND-- #182
*** 3 ballots ***
batch 2703-C-3, # 43
--AND-- #195
*** 2 ballots ***
batch 2703-C-4, #103
--AND-- #104
*** 2 ballots ***
batch 2703-D-1, # 41

Appendix L:
Timing Data Collection Forms

Available at:
Ballot-level comparison and ballot polling -
https://drive.google.com/open?id=1SrG6iGbIGF4OrIfyoPbJRtUu0TyOSDQj

Batch comparison -
https://drive.google.com/open?id=1GpWziE4fLgNjDvT5-a6Q0-YnvGNwDYER

Appendix M: Collected Timing Data

BALLOT RETRIVAL TIME PER BALLOT FOR BALLOT POLLING

	COUNTING		K-CUT		RULER		SCALE	
all data								
sample size (n	5	sample size (	48	sample size (	47	sample size (r	57
avg	2:07	avg	1:32	avg	1:20	avg	0:41
median	1:44	median	1:26	median	1:15	median	0:35
min	0:51	min	0:39	min	0:13	min	0:05
max	3:21	max	5:01	max	3:38	max	2:08
all new box (Y)		all new box (Y)		all new box (Y)		all new box (Y)	
sample size (n	2	sample size (	4	sample size (	4	sample size (r	7
avg	1:33	avg	1:22	avg	1:10	avg	0:55
median	1:33	median	1:23	median	1:04	median	1:05
all New Folder Y		all New Folder Y		all New Folder Y		all New Folder Y	
sample size (n	4	sample size (	21	sample size (	17	sample size (r	25
avg	1:49	avg	1:44	avg	1:17	avg	0:51
median	1:44	median	1:32	median	1:10	median	0:48
Ballot in same folder & box		Ballot in same folder and box		Ballot in same folder and box		Ballot in same folder and box	
sample size (n	1	sample size (	27	sample size (	31	sample size (r	32
avg	3:18	avg	1:22	avg	1:19	avg	0:34
median	3:18	median	1:23	median	1:15	median	0:32

(Based on timing of individual ballot retrieval in minutes)

COUNTING	3:50	K-CUT	1:44	RULER	1:18	SCALE	1:18

(Based on overall timing at table 1 by Tony adams)

BALLOT RETRIEVAL	
all data	
sample size (n	79
avg	0:51
median	0:45
min	0:10
max	4:15
all New box Y	
sample size (n	10
avg	1:07
median	1:01
all new folder (Y)	
sample size (n	39
avg	0:57
median	0:48
Ballot in same folder & box	
sample size (n	22
avg	0:35
median	0:31

BALLOT COMPARISON AUDIT TIMES

Average Time per ballot to evaluate one contest:	0.42 minutes per ballot
Average Time per ballot to evaluate ten contests:	0.98 minutes per ballot

BATCH COMPARISON AUDIT TIMES

Batch Comparison Set-up time per ballot in minutes (Orienting and batching ballots, Steps 0 - 4) with 4 persons per team

sample size	3
avg	0.27
median	0.34
min	0.10
max	0.38

Batch Comparison Tally Evaluation time per ballot in minutes with 4 persons per team (Steps

sample size (n)	12
avg	0:17
median	0:12
min	0:11
max	0:53

Summary for Sort and Stack timing with 2 persons per team

Sort & Stack times min:sec	Sort & Stack times in sec	# of Ballots	Time per Ballot in seconds
18:59	**1139**	127	9
17:35	**1055**	127	8
17:17	1037	127	8
12:18	**738**	110	7
9:34	**574**	110	5
9:50	**590**	110	5
14:19	859	300	3
			46

Average	7
Median	7

Appendix N: Principles and Best Practices for Post-Election Tabulation Audits

Available at:
https://www.verifiedvoting.org/wp-content/uploads/2019/01/Audit-Principles-Best-Practices-2018.pdf#page=7

Appendix O: CVR Data Format Used in Rhode Island

Available at:
https://drive.google.com/file/d/1qMNSbB3tSrLIMe1GBxUWIdJHBZ9uWeC1/view?usp=sharing

Appendix P: Rhode Island Sample Ballot - Portsmouth

Mail Voter

State of Rhode Island Official Ballot
General Election
November 6, 2018
Portsmouth

Precinct 2701
Congressional District 1
Senate District 11
Representative District 69

To mark your choice
Fill in the oval to the left of your choice.
Do not make any marks outside of the oval.

To vote for a write-in candidate
To vote for a candidate whose name is not on the ballot:
Fill in the oval to the left of Write-in and print the name clearly in the box.

Senator in Congress
Six Year Term
Vote for 1
- ○ **Sheldon Whitehouse** DEMOCRAT
- ○ **Robert G. Flanders, Jr.** REPUBLICAN
- ○ **Write-in**

Representative in Congress District 1
Two Year Term
Vote for 1
- ○ **David N. Cicilline** DEMOCRAT
- ○ **Patrick J. Donovan** REPUBLICAN
- ○ **Write-in**

Governor
Four Year Term
Vote for 1
- ○ **Gina M. Raimondo** DEMOCRAT
- ○ **William H. Gilbert** MODERATE
- ○ **Allan W. Fung** REPUBLICAN
- ○ **Anne Armstrong** Compassion
- ○ **Luis Daniel Munoz** Independent
- ○ **Joseph A. Trillo** Independent
- ○ **Write-in**

Lieutenant Governor
Four Year Term
Vote for 1
- ○ **Daniel J. McKee** DEMOCRAT
- ○ **Joel J. Hellmann** MODERATE
- ○ **Paul E. Pence** REPUBLICAN
- ○ **Jonathan J. Riccitelli** Independent
- ○ **Ross K. McCurdy** Independent
- ○ **Write-in**

Secretary of State
Four Year Term
Vote for 1
- ○ **Nellie M. Gorbea** DEMOCRAT
- ○ **Pat V. Cortellessa** REPUBLICAN
- ○ **Write-in**

Attorney General
Four Year Term
Vote for 1
- ○ **Peter F. Neronha** DEMOCRAT
- ○ **Alan Gordon** Compassion
- ○ **Write-in**

General Treasurer
Four Year Term
Vote for 1
- ○ **Seth Magaziner** DEMOCRAT
- ○ **Michael G. Riley** REPUBLICAN
- ○ **Write-in**

Senator in General Assembly District 11
Two Year Term
Vote for 1
- ○ **James Arthur Seveney** DEMOCRAT
- ○ **Stephanie L. Calise** REPUBLICAN
- ○ **Write-in**

Representative in General Assembly District 69
Two Year Term
Vote for 1
- ○ **Susan R. Donovan** DEMOCRAT
- ○ **Douglas W. Gablinske** Independent
- ○ **Write-in**

School Committee
Four Year Term
Vote for any 4
- ○ **Catherine H. Holtman** DEMOCRAT
- ○ **Allen J. Shers** REPUBLICAN
- ○ **Thomas Richard Vadney** REPUBLICAN
- ○ **John Amos Schlesinger** REPUBLICAN
- ○ **Frederick W. Faerber, III** Independent
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**

Town Council
Two Year Term
Vote for any 7
- ○ **Linda L. Ujifusa** DEMOCRAT
- ○ **Paul Francis Kesson** REPUBLICAN
- ○ **Kevin M. Aguiar** DEMOCRAT
- ○ **Jeffrey L. Richard** REPUBLICAN
- ○ **J. Mark Ryan** DEMOCRAT
- ○ **Keith E. Hamilton** REPUBLICAN
- ○ **Andrew V. Kelly** DEMOCRAT
- ○ **Lawrence J. Fitzmorris** REPUBLICAN
- ○ **Raymond Douglas Davis** DEMOCRAT
- ○ **Debra Cardoza** REPUBLICAN
- ○ **Leonard Barry Katzman** DEMOCRAT
- ○ **Elizabeth A. Pedro** REPUBLICAN
- ○ **Daniela T. Abbott** DEMOCRAT
- ○ **David M. Gleason** Independent
- ○ **Peter D. Roberts** Independent
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**

State Questions On Back

State Questions

QUESTIONS 1 - 3
(Chapter 047 - Public Laws 2018)

Shall the action of the General Assembly, by an act passed at the January 2018 session, authorizing the issuance of bonds, refunding bonds, and/or temporary notes of the State of Rhode Island for the capital projects and in the amount with respect to each such project listed below (Questions 1-3) be approved, and the issuance of bonds, refunding bonds, and/or temporary notes authorized in accordance with the provisions of said act?

1. RHODE ISLAND SCHOOL BUILDINGS - $250,000,000

To provide state assistance to cities and towns for the construction of new public schools and renovation of existing public schools.
- ○ **Approve**
- ○ **Reject**

2. HIGHER EDUCATION FACILITIES - $70,000,000

For higher education facilities, to be allocated as follows:

(a) University of Rhode Island Narragansett Bay Campus - $45,000,000

(b) Rhode Island College School of Education and Human Development - $25,000,000
- ○ **Approve**
- ○ **Reject**

3. GREEN ECONOMY AND CLEAN WATER - $47,300,000

For environmental and recreational purposes, to be allocated as follows:

(a) Coastal Resiliency and Public Access Projects - $5,000,000

(b) Capital for Clean Water and Drinking Water - $7,900,000

(c) Wastewater Treatment Facility Resilience Improvements - $5,000,000

(d) Dam Safety - $4,400,000

(e) Dredging - Downtown Providence Rivers - $7,000,000

(f) State Bikeway Development Program - $5,000,000

(g) Brownfield Remediation and Economic Development - $4,000,000

(h) Local Recreation Projects - $5,000,000

(i) Access to Farmland - $2,000,000

(j) Local Open Space - $2,000,000
- ○ **Approve**
- ○ **Reject**

Appendix P: Rhode Island Sample Ballot - Bristol

Mail Voter

State of Rhode Island Official Ballot
General Election
November 6, 2018
Bristol

Precinct 0201
Congressional District 1
Senate District 10
Representative District 68

To mark your choice
Fill in the oval to the left of your choice.
Do not make any marks outside of the oval.

To vote for a write-in candidate
To vote for a candidate whose name is not on the ballot:
Fill in the oval to the left of Write-in and print the name clearly in the box.

Senator in Congress
Six Year Term
Vote for 1
- Sheldon Whitehouse, DEMOCRAT
- Robert G. Flanders, Jr., REPUBLICAN
- Write-in

Representative in Congress District 1
Two Year Term
Vote for 1
- David N. Cicilline, DEMOCRAT
- Patrick J. Donovan, REPUBLICAN
- Write-in

Governor
Four Year Term
Vote for 1
- Gina M. Raimondo, DEMOCRAT
- William H. Gilbert, MODERATE
- Allan W. Fung, REPUBLICAN
- Anne Armstrong, Compassion
- Luis Daniel Munoz, Independent
- Joseph A. Trillo, Independent
- Write-in

Lieutenant Governor
Four Year Term
Vote for 1
- Daniel J. McKee, DEMOCRAT
- Joel J. Hellmann, MODERATE
- Paul E. Pence, REPUBLICAN
- Jonathan J. Riccitelli, Independent
- Ross K. McCurdy, Independent
- Write-in

Secretary of State
Four Year Term
Vote for 1
- Nellie M. Gorbea, DEMOCRAT
- Pat V. Cortellessa, REPUBLICAN
- Write-in

Attorney General
Four Year Term
Vote for 1
- Peter F. Neronha, DEMOCRAT
- Alan Gordon, Compassion
- Write-in

General Treasurer
Four Year Term
Vote for 1
- Seth Magaziner, DEMOCRAT
- Michael G. Riley, REPUBLICAN
- Write-in

Senator in General Assembly District 10
Two Year Term
Vote for 1
- Walter S. Felag, Jr., DEMOCRAT
- Write-in

Representative in General Assembly District 68
Two Year Term
Vote for 1
- Laufton Ascencao, DEMOCRAT
- William James Hunt, Jr., Libertarian
- Write-in

Town Administrator
Two Year Term
Vote for 1
- Steven Contente, Independent
- Write-in

Town Clerk
Two Year Term
Vote for 1
- Melissa Cordeiro, DEMOCRAT
- Louis P. Cirillo, REPUBLICAN
- Write-in

Town Council
Two Year Term
Vote for any 5
- Nathan T. Calouro, DEMOCRAT
- Thomas M. Carroll, REPUBLICAN
- Aaron J. Ley, DEMOCRAT
- William L. Sousa Grapentine, REPUBLICAN
- Carolyn P. Medina, DEMOCRAT
- Patrick M. McCarthy, REPUBLICAN
- Timothy Edward Sweeney, DEMOCRAT
- Mary A. Parella, REPUBLICAN
- Ethan M. Tucker, DEMOCRAT
- Antonio A. Teixeira, Independent
- Write-in
- Write-in
- Write-in
- Write-in
- Write-in

Non-Partisan Regional School Committee
Four Year Term
Vote for any 3
- Andrew Benn
- Carly Reich
- William M. O'Dell
- Robert D. Hancock
- Patrick Usher, Jr.
- Victor G. Cabral, Jr.
- Diane B. Campbell
- Paul Silva
- Sheila O. Ellsworth
- Write-in
- Write-in
- Write-in

State Questions On Back

State Questions

QUESTIONS 1 - 3
(Chapter 047 - Public Laws 2018)

Shall the action of the General Assembly, by an act passed at the January 2018 session, authorizing the issuance of bonds, refunding bonds, and/or temporary notes of the State of Rhode Island for the capital projects and in the amount with respect to each such project listed below (Questions 1-3) be approved, and the issuance of bonds, refunding bonds, and/or temporary notes authorized in accordance with the provisions of said act?

1. RHODE ISLAND SCHOOL BUILDINGS - $250,000,000

To provide state assistance to cities and towns for the construction of new public schools and renovation of existing public schools.
- Approve
- Reject

2. HIGHER EDUCATION FACILITIES - $70,000,000

For higher education facilities, to be allocated as follows:

(a) University of Rhode Island Narragansett Bay Campus - $45,000,000

(b) Rhode Island College School of Education and Human Development - $25,000,000
- Approve
- Reject

3. GREEN ECONOMY AND CLEAN WATER - $47,300,000

For environmental and recreational purposes, to be allocated as follows:

(a) Coastal Resiliency and Public Access Projects - $5,000,000

(b) Capital for Clean Water and Drinking Water - $7,900,000

(c) Wastewater Treatment Facility Resilience Improvements - $5,000,000

(d) Dam Safety - $4,400,000

(e) Dredging - Downtown Providence Rivers - $7,000,000

(f) State Bikeway Development Program - $5,000,000

(g) Brownfield Remediation and Economic Development - $4,000,000

(h) Local Recreation Projects - $5,000,000

(i) Access to Farmland - $2,000,000

(j) Local Open Space - $2,000,000
- Approve
- Reject

Appendix P: Rhode Island Sample Ballot - Cranston

Mail Voter

State of Rhode Island Official Ballot
General Election
November 6, 2018
Cranston

Precinct 0704
Congressional District 2
Senate District 28
Representative District 14
Cranston Ward 2

To mark your choice
Fill in the oval to the left of your choice.
Do not make any marks outside of the oval.

or write-in:

To vote for a write-in candidate
To vote for a candidate whose name is not on the ballot:
Fill in the oval to the left of Write-in and print the name clearly in the box.

Senator in Congress
Six Year Term
Vote for 1

- ○ **Sheldon Whitehouse** DEMOCRAT
- ○ **Robert G. Flanders, Jr.** REPUBLICAN
- ○ **Write-in**

Representative in Congress District 2
Two Year Term
Vote for 1

- ○ **James R. Langevin** DEMOCRAT
- ○ **Salvatore G. Caiozzo** REPUBLICAN
- ○ **Write-in**

Governor
Four Year Term
Vote for 1

- ○ **Gina M. Raimondo** DEMOCRAT
- ○ **William H. Gilbert** MODERATE
- ○ **Allan W. Fung** REPUBLICAN
- ○ **Anne Armstrong** Compassion
- ○ **Luis Daniel Munoz** Independent
- ○ **Joseph A. Trillo** Independent
- ○ **Write-in**

Lieutenant Governor
Four Year Term
Vote for 1

- ○ **Daniel J. McKee** DEMOCRAT
- ○ **Joel J. Hellmann** MODERATE
- ○ **Paul E. Pence** REPUBLICAN
- ○ **Jonathan J. Riccitelli** Independent
- ○ **Ross K. McCurdy** Independent
- ○ **Write-in**

Secretary of State
Four Year Term
Vote for 1

- ○ **Nellie M. Gorbea** DEMOCRAT
- ○ **Pat V. Cortellessa** REPUBLICAN
- ○ **Write-in**

Attorney General
Four Year Term
Vote for 1

- ○ **Peter F. Neronha** DEMOCRAT
- ○ **Alan Gordon** Compassion
- ○ **Write-in**

General Treasurer
Four Year Term
Vote for 1

- ○ **Seth Magaziner** DEMOCRAT
- ○ **Michael G. Riley** REPUBLICAN
- ○ **Write-in**

Senator in General Assembly District 28
Two Year Term
Vote for 1

- ○ **Joshua Miller** DEMOCRAT
- ○ **Write-in**

Representative in General Assembly District 14
Two Year Term
Vote for 1

- ○ **Charlene Lima** DEMOCRAT
- ○ **Write-in**

Council - City Wide
Two Year Term
Vote for any 3

- ○ **Steven A. Stycos** DEMOCRAT
- ○ **Michael J. Farina** REPUBLICAN
- ○ **Gail D. Harvey** DEMOCRAT
- ○ **Kenneth J. Hopkins** REPUBLICAN
- ○ **Sarah K. Lee** DEMOCRAT
- ○ **Amy L. Ricci** REPUBLICAN
- ○ **Write-in**
- ○ **Write-in**
- ○ **Write-in**

Council Ward 2
Two Year Term
Vote for 1

- ○ **Paul J. McAuley** DEMOCRAT
- ○ **Write-in**

Non-Partisan School Committee - City Wide
Two Year Term
Vote for 1

- ○ **Michael Anthony Traficante**
- ○ **Write-in**

Non-Partisan School Committee Ward 2
Two Year Term
Vote for 1

- ○ **Stephanie Giangrande Culhane**
- ○ **Write-in**

State and Local Questions On Back

State Questions

QUESTIONS 1 - 3
(Chapter 047 - Public Laws 2018)

Shall the action of the General Assembly, by an act passed at the January 2018 session, authorizing the issuance of bonds, refunding bonds, and/or temporary notes of the State of Rhode Island for the capital projects and in the amount with respect to each such project listed below (Questions 1-3) be approved, and the issuance of bonds, refunding bonds, and/or temporary notes authorized in accordance with the provisions of said act?

1. RHODE ISLAND SCHOOL BUILDINGS - $250,000,000

To provide state assistance to cities and towns for the construction of new public schools and renovation of existing public schools.

- ○ **Approve**
- ○ **Reject**

2. HIGHER EDUCATION FACILITIES - $70,000,000

For higher education facilities, to be allocated as follows:

(a) University of Rhode Island Narragansett Bay Campus - $45,000,000

(b) Rhode Island College School of Education and Human Development - $25,000,000

- ○ **Approve**
- ○ **Reject**

3. GREEN ECONOMY AND CLEAN WATER - $47,300,000

For environmental and recreational purposes, to be allocated as follows:

(a) Coastal Resiliency and Public Access Projects - $5,000,000

(b) Capital for Clean Water and Drinking Water - $7,900,000

(c) Wastewater Treatment Facility Resilience Improvements - $5,000,000

(d) Dam Safety - $4,400,000

(e) Dredging - Downtown Providence Rivers - $7,000,000

(f) State Bikeway Development Program - $5,000,000

(g) Brownfield Remediation and Economic Development - $4,000,000

(h) Local Recreation Projects - $5,000,000

(i) Access to Farmland - $2,000,000

(j) Local Open Space - $2,000,000

- ○ **Approve**
- ○ **Reject**

Local Question

4. PLAYGROUNDS AND ATHLETIC FIELDS $2,000,000 BONDS AND NOTES
(Ordinance of the City Council adopted June 25, 2018)

Shall Two Million Dollars ($2,000,000) debt of the City be incurred for the purpose of financing the acquisition, construction, improvement, renovation, repair, alteration and equipping of playgrounds and athletic fields in the City under Rhode Island General Laws § 45-12-2 and an order of the City Council that became effective on June 25, 2018?

- ○ **Approve**
- ○ **Reject**

Cranston

Common Cause Rhode Island

verifiedvoting

BRENNAN CENTER FOR JUSTICE