A Risk-Limiting Audit In Denmark
A Pilot

Carsten Schürmann *

DemTech
IT University of Copenhagen
carsten@itu.dk

Abstract. The theory of risk-limiting audits is well-understood, at least mathematically. Such audits serve to create confidence in the reported election outcome by checking the evidence created during the election. When election officials introduce election technologies into the voting process, it is best to do this after the appropriate auditing framework has been implemented. In this paper, we describe our experiences with piloting a risk-limiting audit of a referendum that was held in Denmark on December 3, 2015. At the time of the publication of this paper, Denmark's election law did not permit electronic voting technologies to be used during voting allowing us to study auditing in isolation.
Our findings are that (1) risk-limiting audits also apply to paper and pencil elections; (2) election officials usually support risk-limiting audits even if no voting technologies are used because these audits can improve the efficiency of the manual count; (3) that practical and organizational challenges must be overcome to keep audits repeatable, in particular it must be possible to identify individual ballots repeatedly and reliably; (4) it is possible to arrange an audit for the result of an earlier stage in a count during a later stage, for example, an audit of the rough count results fine count; and (5) that whenever the electronic voting technologies are considered, auditing should be considered as part of feasibility study.

1 Introduction

A voter verifiable paper audit trail is only as good as its curation and the auditing procedure that uses it to check the validity of the election result. If a voting machine creates a paper audit trail, it is only meaningful when it is inspected in a valid and systematic fashion, for example, by means of a manual recount or a risk-limiting audit. Electronic voting technologies should therefore only be used *after* a suitable auditing framework and the relevant auditing procedures have been defined and implemented.

* This publication was made possible in part by the DemTech grant 10-092309 from the Danish Council for Strategic Research, Program Commission on Strategic Growth Technologies and in part by NPRP Grant #7-988-1-178 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

In 2013, Denmark held an European parliament election, for which we attempted to pilot a first nation-wide risk limiting audit across jurisdictional boundaries with the goal to learn about the requirements and challenges to Danish election processes and procedures. Unfortunately, the pilot failed, because too few constituencies participated in this pilot where participation was voluntary. Audits require that ballot papers are stored in such a way that they can be easily retrieved. This can be challenging, especially for larger constituencies that have to curate large stacks of ballot papers, draw large random samples, and have only limited resources to conduct the audit. The Copenhagen constituency is so big, for example, that they rent extra containers to store all ballot papers. This made us wonder *what are the main challenges for adopting risk-limiting audits in Denmark, and how can election procedures be modified to render them feasible.* In 2013, Denmark held a referendum, where we expected large margins, and so we took a second attempt to undertake a national risk-limiting audit, but this time we only focused only on Copenhagen.

At the time of the publication of this paper, Denmark's election law did not permit electronic voting technologies to be used during voting. None the less, studying auditing in the setting of paper-only based elections is important to understand the implications of auditing frameworks on the electoral process in preparation, for example, for nation states that consider introducing electronic voting technologies. Risk-limiting audits allow auditors to inspect a small sample of ballots to assess the overall quality of the election result. Risk-limiting audits have been developed by Stark and Lindeman [LS12] in the context of US elections, where the use of electronic voting machines with voter verifiable paper audit trails is prevalent. Stark and Lindeman ask, how to audit the election result computed by an electronic vote tabulation system against the paper trail. For our pilot, we audited the election result against manually completed ballot papers and not a computer generated paper audit trail. The hallmark characteristic of our audit is that we audited the election night results during the fine count phase of the election the very next day. In Denmark, votes are counted multiple times. The *rough count* takes place on election night and the results are usually published before midnight. The next morning, the *fine count* determines the official election result. In our pilot, we audited the result of the rough count during the fine count.

We believe that the findings that we describe in this paper provide valuable insights that are applicable beyond the Danish context and might be of interest to any election commission that considers integrating an auditing framework into their respective administrative processes, for example, as a supplementary quality control mechanisms or in preparation for the introduction of electronic voting technologies into their respective electoral processes. The findings pertain to adjusting the legal framework, devising organizational rules with the goal to create confidence among the electorate. This paper is organized as follows. We first revisit the theory of risk-limiting audits in Section 2, then we discuss the design and execution of the audit applied the 2015 referendum on in Section 3, finally we discuss our findings in Section 4 before concluding in Section 5.

2 Risk-Limiting Audits

A risk-limiting audit is a statistical method to create confidence in the correctness of an election result by checking samples of paper ballots. Lindeman and Stark [LS12] distinguish *ballot-polling audits*, where they draw a carefully chosen random sample of ballots to check whether the sample gives sufficiently strong evidence for the correctness of the published election result. In contrast, a *comparison audit* checks the ballot interpretation for a random sample during the audit against their respective interpretation in a vote tabulation system.

One of the requirements is that a ballot manifest is available that describes in detail how ballots are organized and stored, including how many stacks there are and how many ballots can be found in each stack. This information establishes a total order among all ballots that is needed for drawing the sample and being able to retrieve individual ballots for inspection. For this pilot, we have chosen a comparison audit, which follows the steps defined in [LS12] and described below and commences only after the election has closed and the election results were published.

1. Entropy is generated, for example by throwing physical dice. The entropy is subsequently used to initialize the random generator. We denote entropy with e.
2. From the election result we compute the smallest margin of the election. There are many different margins to consider, we only focus on one-vote overstatements where ballots were accidentally not counted for a loosing candiate, because they were either invalid, or counted to someone else but the winner and two-vote overstatements, which are ballots that were erroneously counted for the winning candidate instead of the loosing candidate. For more details on margin computation, consult [Sta10].
3. Using some more statistics [Sta10], we determine the number of ballots to be audited. We denote this number with k. Note, that one input to this computation is the total number of ballots cast in the contest to which we refer with n.
4. Using the algorithm depicted in Figure 1, we compute a set of ballots to be audited, the random sample, from inputs e, k, n. As the audit needs to be reproducible, we cannot simply apply the random generator provided by an implementation language. Instead, we trust that hashing strings gives us a uniform distribution. In the implementation is done using `sha256` (see line 6. in Figure 1). The result of the algorithm is a list of random numbers, between the 0 and $n-1$, each identifies a particular ballot in the total order. Because the computation is parametric only in e, k, and n, the result is reproducible.
5. Using the total order given by the ballot manifest, we can now compute the precise location (in terms of municipality, polling station, party name, candidate name, for example) for each ballot to be audited.

Table 1. Drawing a random sample of ballots

```
 1: function drawSample(string e; int k; int n)
 2:     i = 1
 3:     bs = {}
 4:     while i ≤ k do do
 5:         x = e + "," + intToString(i)
 6:         y = sha256(x)
 7:         z = decimal(y)
 8:         b = z mod (n − 1)
 9:         bs = bs ∪ {b}
10:     end while
11:     return(bs)
12: end function
```

3 Referendum

We conducted a risk-limiting audit for a national referendum in Denmark that took place on December 3, 2015[1]. Elections in Denmark allow for advance voting, where voters can vote form anywhere in the country (and abroad), but their votes will be registered in their respective constituencies. Different from other countries, that allow voters to complete their ballots at home and mail them in, Denmark requires voters to cast their vote in controlled environments using the double envelope technique; officials will then mail the envelopes to the voter's behalf to their respective polling places, where they are registered on the electoral roll before the polling places open in the morning of the election day.

Denmark consists of 92 constituencies, where Copenhagen and Aarhus are the largest. Copenhagen alone organizes 50 polling stations.

Denmark does not offer computer-enabled vote casting. It neither offers internet voting as an alternate voting channel, nor electronic ballot markers. Special rules apply to voters with disabilities, who may take a friend to help complete a ballot but are required to have an election official present to witness that the voter's intent is correctly reflected on the ballot. All ballots are completed by pencil and paper and then hand-counted by elections officials in the evening of election day, which is called *rough count*, and again during the next day which is called *fine count*. This means, in Denmark, ballots are counted several times.

Computers, however, play an important part in other parts of the electoral process. The electoral roll is digitally copied from the public registry of all residents of Denmark a few days prior to the election. Every voter receives a document (*valgkort*) in the mail, allowing them to register to vote on election day. For municipalities that support digital voter lists this document contains also a computer readable barcode. The barcode is read and the voter is then digitally

[1] `http://www.dst.dk/valg/Valg1664255/valgopg/valgopgHL.htm`

Table 2. Official results of the 2015 Referendum

YES	1.375.862	46,9%
NO	1.558.437	53,1%
Total number of valid votes	2.934.299	
Blank votes	48.216	
Other invalid votes	7.746	
Total number of invalid votes	55.962	
Total number of votes	2.990.261	

checked off the electoral role. Other municipalities check voters manually of the electoral role.

During election night, computers are used to report the results from each polling station to the ministry. To verify that the reporting was ok, the ministry expects a telephone call from each polling station within five minutes of the original reporting, where subsequently all totals are manually checked. The final seat assignment for national parliament, European parliament, or municipal elections is usually computed by a computer program in the ministry using a combination of D'Hondt and Saint Leguë methods. For the referendum that we describe here, the computer program only added the totals from each reporting constituency. The results of the referendum contest are depicted in Figure 2. The referendum was rejected by majority.

3.1 Designing the Audit

The audit we have chosen was a comparison ballot audit [LS12], where we compare the interpretation of ballot in the random sample to how it was recorded. We designed our audit in such a way that we would audit the result of the rough count (on election night) during the fine count, which took place the next day. Fortunate for us, the company that conducts the election night reporting, publishes all results online. Scraping their webpage allowed us to compute the ballot manifest, which we subsequently used to draw the random sample.

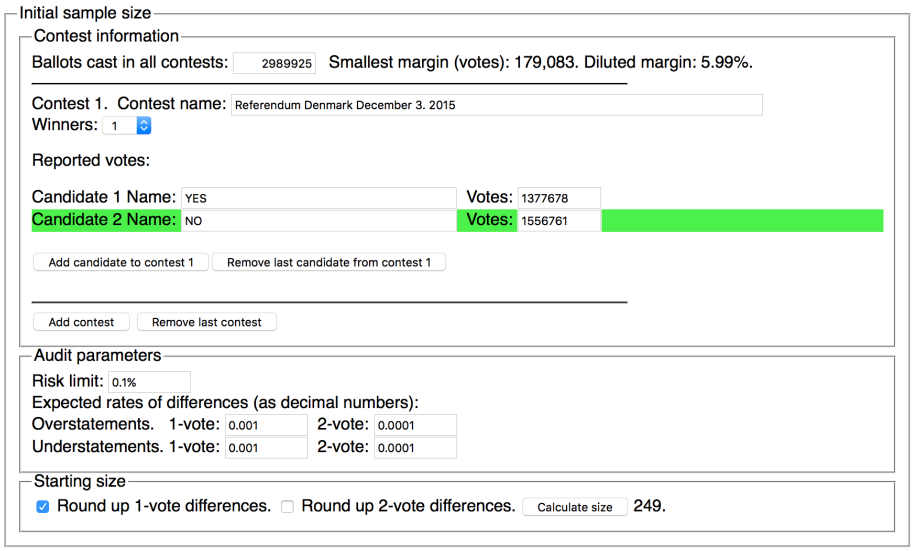
This referendum is the kind of election [Sta10] can be applied to directly. We decided for a risk limit of 0.1%, which means that after the audit, we can be 99.9% sure that the published election outcome is correct. Note that the following rough count totals differ from the official elections result described in Figure 2.

YES	1.377.678
NO	1.556.761
Total number of votes	2.989.925

As these were the only numbers we had, we used them to compute the margin for our risk limiting audit, using Stark's online resources[2]. A screenshot of our

[2] `http://www.stat.berkeley.edu/~stark/Vote/auditTools.htm#`

Table 3. Screenshot of Stark's Risk Limiting Audit Tool



interaction with the tool is depicted in Figure 3, with the result, that the margin is 179083 votes, the diluted margin is 5.99% and the number of ballots to audit is 249. As the margin is large, it is perhaps not surprising that the sample size is relatively small.

We derived entropy by rolling the dice and so we could run the algorithm from Figure 1 with arguments

$$
\begin{aligned}
e &= \texttt{"674987539957481874"} \\
k &= 249 \\
n &= 2.989.925
\end{aligned}
$$

to draw the sample. Out of the 249 ballots, 36 fell into the Copenhagen constituency, the other 213 were distributed among the rest of Denmark. We focused the audit exclusively on Copenhagen, because (1) this was only a pilot, where we wanted to learn about the mechanisms behind executing the audit, and not really audit the correctness of the election result. It is worth mentioning that we found no misinterpretation of any of the 36 ballots that we checked. (2) Copenhagen is the biggest municipality in Denmark, which means, that if we can have a successful audit in Copenhagen, then risk-limiting audits could be – at least in principle – be executed anywhere in the country. The list of all ballots that we audited can be found in Figure 4.

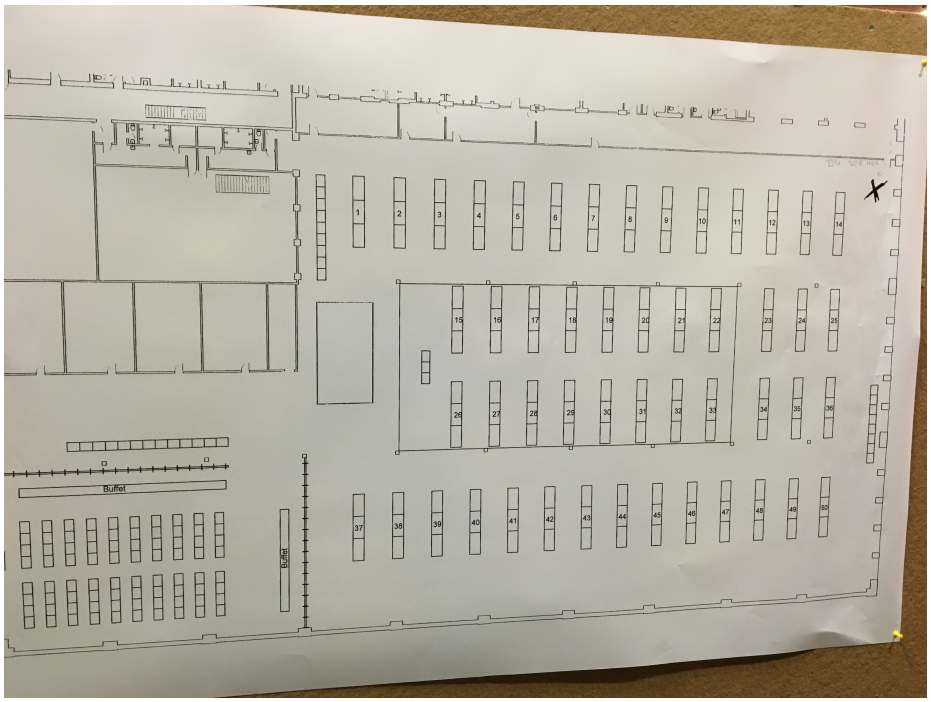
3.2 Executing the Audit

With the sample computed, we, a team of academic election researchers, went to the fine count center for the constituency of Copenhagen on the morning

Table 4. List of Ballots to be Audited

Bispebjergkredsen	6.	Syd	Ballot 2037 interpreted as NO?
Bispebjergkredsen	6.	Øst	Ballot 650 interpreted as YES?
Brønshøjkredsen	7.	Kirkebjerg	Ballot 821 interpreted as YES?
Brønshøjkredsen	7.	Nordvest	Ballot 864 interpreted as YES?
Brønshøjkredsen	7.	Syd	Ballot 1624 interpreted as NO?
Indre Bykredsen	3.	Indre By	Ballot 2305 interpreted as YES?
Indre Bykredsen	3.	Indre By	Ballot 4434 interpreted as YES?
Indre Bykredsen	3.	Syd	Ballot 2243 interpreted as YES?
Indre Bykredsen	3.	Øst	Ballot 1750 interpreted as YES?
Indre Bykredsen	3.	Øst	Ballot 4926 interpreted as YES?
Indre Bykredsen	3.	Øst	Ballot 2390 interpreted as NO?
Nørrebrokredsen	5.	Nordvest	Ballot 1925 interpreted as YES?
Nørrebrokredsen	5.	Nordvest	Ballot 1124 interpreted as NO?
Nørrebrokredsen	5.	Nordvest	Ballot 1238 interpreted as NO?
Nørrebrokredsen	5.	Vest	Ballot 463 interpreted as NO?
Sundbyvesterkredsen	2.	Nord	Ballot 3496 interpreted as YES?
Sundbyvesterkredsen	2.	Nord	Ballot 1890 interpreted as NO?
Sundbyvesterkredsen	2.	Nord	Ballot 2630 interpreted as NO?
Sundbyvesterkredsen	2.	Nord	Ballot 4275 interpreted as NO?
Sundbyvesterkredsen	2.	Sundbyvester	Ballot 2965 interpreted as YES?
Sundbyvesterkredsen	2.	Sundbyvester	Ballot 3994 interpreted as NO?
Sundbyvesterkredsen	2.	Syd	Ballot 2019 interpreted as YES?
Sundbyvesterkredsen	2.	Vest	Ballot 2584 interpreted as NO?
Sundbyøsterkredsen	4.	Nord	Ballot 1864 interpreted as NO?
Sundbyøsterkredsen	4.	Nord	Ballot 4917 interpreted as NO?
Sundbyøsterkredsen	4.	Syd	Ballot 2080 interpreted as YES?
Valbykredsen	8.	Midt	Ballot 644 interpreted as YES?
Valbykredsen	8.	Nord	Ballot 346 interpreted as YES?
Valbykredsen	8.	Sydøst	Ballot 260 interpreted as YES?
Valbykredsen	8.	Valby	Ballot 2469 interpreted as YES?
Valbykredsen	8.	Valby	Ballot 2642 interpreted as YES?
Vesterbrokredsen	9.	Nord	Ballot 2533 interpreted as YES?
Østerbrokredsen	1.	Nord	Ballot 196 interpreted as YES?
Østerbrokredsen	1.	Syd	Ballot 5465 interpreted as YES?
Østerbrokredsen	1.	Østerbro	Ballot 2706 interpreted as NO?
Østerbrokredsen	1.	Østerbro	Ballot 3476 interpreted as NO?

Table 5. Layout of the Fine Count Location



after the election. The night before, the ballots from each polling center were transported (by taxi) to the fine counting center under constant surveillance of the election officials who would spend the night there. The fine count commences in the following way. The ballot papers from each polling station are placed on a separate table and counted by civil servants. Each table is numbered. Once the fine count is complete, the results are being entered into a database, and the official in charge of the fine count (*rådhusbetjent*) will forward a piece of paper with the number of table that completed to recount to the quality control office. The paper is important, whoever has the paper, has the right to work on the results of the respective table. During quality control, a team will assess the probability, using statistical tools, if errors were made. In the case that the probability is low, the table is officially done, the piece of paper will be returned to the official in charge of the find count, and the ballots will eventually be packed for further storage. In the case that the probability is deemed unacceptable, civil servants return to the table to recount and find the mistake.

The map of the layout of the different tables can be seen on the map depicted in Figure 5. The large white square between table 1 and table 37 is the table of the official in charge of the fine count. The quality control office is to the left of this table and not clearly marked.

The head of elections of Copenhagen municipality gave us permission to conduct the audit. We were permitted to touch the ballot papers of any table that has passed quality control and awaited packing. In preparation for the audit, we determined together with the head of elections, which of tables contained the relevant ballots of our sample described in Figure 4. The person in charge of the fine count then kept a stack of the pieces of paper identifying the tables scheduled for auditing and passed them along to us after the table cleared quality control.

On each table, we approached, ballots were packed with rubber bands into stacks of 100, and sometimes 5 such stacks were made into a larger stack, also with rubber bands. On some tables, the stacks were already packed into boxes inducing some kind of order, on other tables, the stacks were placed in no particular order. Advanced votes (which look different in Denmark, because the official ballot form is not yet approved when advanced voting commences) were usually rubber banded and all stacks of a 100 were slipped into one large envelope. It was up to us, the auditors, to fix the order of the ballots.

As depicted in Figure 4, some ballot numbers were rather high, 4275 being the highest. From the outset, we decided to assume that all packs of 100 contained a 100 votes, so that finding the correct ballot became feasible in the time frame provided. Also we would consider election day ballots always before considering mail-in ballots. For ballot 4275, for example, we would select the 43. pack of 100 and then start counting backwards until we arrive at ballot 75. If we could not discern a clear order, we would define one; but because we wer handling live ballots, we could not document the order that we have imposed during our audit.

4 Findings

4.1 Paper Elections

Comparison audits do not only apply to elections with an electronic ballot interpretation — they also apply to paper-based elections. We only have to presume that we know the interpretation of each ballot for manual vote tabulation. Our pilot is to our knowledge the first application of a comparison audit to a paper-only election and it shows that it is possible to incorporate the audit into the counting process. A consequence is that countries that plan to transition from pen and paper to electronic elections, can already set up the auditing framework and related processes before even the feasibility study and procurement of election technologies commences.

4.2 Election Officials' Reaction

In our experience, the term *risk-limiting audits* does not sit well with election officials who are not used to the idea of post-election auditing. A few election boards that we interacted with in the past were quite offended by this term. Judging from their reactions, they must have perceived this a sign of distrust

that they were informed that their work shall undergo another round of scrutiny. We suspect that a better term is *result-confirming* audits, which is more in line with the interest of most election officials. On the other hand, our pilot was also met with a lot enthusiasm, in particular, it was felt that risk-limiting audits may actually speed up the quality checking and error finding during the fine count. Although this might be true, our pilot shows that there was a noticeable difference between the election results reported after rough count and fine count, respectively, and this difference can only be corrected by a full recount.

4.3 Repeatability

An audit is only then meaningful, if it can be retraced by an independent party of auditors auditing the auditors. In this regards, we noticed, that it is by no means clear if the same set of ballots that we audited could at least in theory be retrieved again. After we left each table, a team of packers arrived, packing the ballots carefully according to the rules, which are clearly specified in Danish election law (*valgloven*). The order we committed to during the audit was likely destroyed when packing the ballots. We conclude that any election that is designed with an audit in mind shall also provide a mechanism for maintaining a particular order for all ballots of each polling station. This order must be preserved when the ballots are stored.

4.4 Integration

A further observation is that of integration of a comparison audit into the larger counting process. We have shown, that it is possible to conduct a risk-limiting audit already one day after the rough count by integrating it into the counting procedures during the fine count. This compares favorably to post-election audits, such as the one conducted in Colorado in 2013 [McB]. This insight suggests that auditing can be an integral part of counting procedures, but it should thought into the process already during the feasibility study.

4.5 Completeness

Our pilot study was not designed as a full scale risk-limiting audit of the Danish referendum. It was more a pilot to understand the changes to the process that occur when conducting such an audit. Were we to scale the audit to all of Denmark, we would need auditors in each fine counting center of which there are 92. Each auditor needs to be properly trained regarding purpose and technicalities. In Copenhagen, counting teams were assigned to the individual tables. Once they finished they were free to leave. We believe it is possible that these teams could also be trained to conduct the audit. And again, such training plans need to be thought carefully into the administrative procedures.

5 Conclusion

In this paper we describe a pilot study for a risk-limiting audit of the 2015 Danish National Referendum. The lessons learned may be of interest to any election managment body that is planning on integrating an auditing framework into their respective electoral processes or that is seeking to introduce electronic voting technology. We conclude that it is advisable to think of auditing as an integral part ot the counting activities and to establish clear rules, procedures, and processes for conducting the audit. In the case that the audit is introduced in conjunction with new election technologies, it is best to consider auditing already as part of the feasibility study. The way that paper evidence and ballots are stored is central for the success and the repeatability of the audit. In the case of internet elections, there is no paper trail, and thus standard auditing techniques do not apply.

References

[LS12] Mark Lindeman and Philip B. Stark. A gentle introduction to risk-limiting audits. *IEEE Security & Privacy*, 10(5):42–49, 2012.

[McB] Neal McBurnett. State of colorado, risk-limiting audit final report, post-election audit initiative grant no. eac110150e. `http://bcn.boulder.co.us/~neal/elections/corla/Risk-Limiting_Audit_Report-Final_20140331.pdf`.

[Sta10] Philip B. Stark. Super-simple simultaneous single-ballot risk-limiting audits. In *Conference on Electronic Voting Technology/Workshop on Trustworthy Elections (EVT/WOTE)*, pages 1–16, 2010.