USENIX Journal of Election Technology and Systems (JETS)
Volume 3, Number 1 • December 2014
www.usenix.org/jets/issues/0301

Verifiable European Elections: Risk-limiting Audits for D'Hondt and its relatives

Philip B. Stark
UC Berkeley
stark@stat.berkeley.edu

Vanessa Teague
University of Melbourne
vjteague@unimelb.edu.au

December 5, 2014

Abstract

We provide Risk Limiting Audits for proportional representation election systems such as D'Hondt and Sainte-Laguë. These techniques could be used to produce evidence of correct (electronic) election outcomes in Denmark, Luxembourg, Estonia, Norway, and many other countries.

1 Introduction

Electronic voting in Europe is both controversial and limited. Some countries use polling-place DREs (direct-recording electronic voting machines); others, such as the Netherlands, Ireland, and Germany, introduced and then rejected DREs. No European country requires auditing a paper trail. Some, including Switzerland, Estonia and (until recently) Norway, use Internet voting systems without universally verifiable tallying.

Risk-limiting audits test an announced election result against voter-verified paper records. They aim to answer the question, "Given an agreed list of cast votes, how do we provide convincing public evidence that the election outcome is correct?"[1] The techniques were developed for plurality voting systems. It is not obvious how to adapt them to complex European election systems. This paper fills the gap for many "highest-averages"[2] proportional representation schemes used in Europe, including D'Hondt and Sainte-Laguë. As far as we know, this is the first work to develop risk-limiting audits for highest-averages proportional representation methods.

We provide several RLA techniques for highest-averages elections: If the reported seat allocation is wrong, there is a guaranteed minimum probability that the audit will correct it.

These methods could be used in Norway, Germany, Luxembourg, Estonia, Denmark, Belgium, and other countries. Our work could apply in Belgium, where—after computer scientists pressured the government—the

[1] Ensuring that the list accurately reflects the voters' intentions is also challenging; see, for instance, Stark and Wagner [2012].

[2] This terminology is from Gallagher [1992].

USENIX Journal of Election Technology and Systems (JETS)
Volume 3, Number 1 • December 2014
www.usenix.org/jets/issues/0301

electronic voting machines produce a paper trail, which has never been audited, despite recommendations [BeV, 2007].[3]

1.1 Background and contribution: Risk-limiting audits

We assume we have a voter-verified paper record that has been determined by a *compliance audit* Benaloh, Jones, Lazarus, Lindeman, and Stark [2011], Lindeman and Stark [2012], Stark and Wagner [2012] to reflect the true electoral outcome (The electoral outcome is the number of seats assigned to each party, not the specific number of votes cast for each party.) We also have a *reported* (electronic) outcome, which we distrust.

Risk-limiting audits (RLAs), introduced by Stark [2008a], provide a statistical assurance that the reported outcome matches the *actual* outcome a full hand tally of the paper record would show. If the reported outcome is wrong, no matter why, a risk-limiting audit has a large probability of correcting it. After a RLA, either there is strong statistical evidence that the outcome is correct, or the outcome is known to be correct.

RLAs have been derived for and performed on plurality contests, majority contests, multi-winner contests, and multiple contests simultaneously [Stark, 2008b, Hall, Miratrix, Stark, Briones, Ginnold, Oakley, Peaden, Pellerin, Stanionis, and Webber, 2009, **?**]. Sarwate, Checkoway, and Shacham [2013] consider risk-limiting audits for IRV/STV, Condorcet and Borda. We know of no work on RLAs for highest-averages systems.

We focus on two approaches to RLAs, described by Lindeman and Stark [2012]: *ballot-polling audits*, which rely on the paper ballots but not the electronic record, and *ballot-level comparison audits*, which compare electronic *cast vote records* (tallies for individual ballots) to the corresponding paper records. Both require a *ballot manifest* that describes how ballots are stored. Ballot-polling audits have minimal set-up costs and need nothing from the electronic system except a reported outcome. But they generally involve inspecting more ballots than ballot-level comparison audits, which require that the voting system report results for individual ballots in a way that allows each to be matched to its corresponding paper record—and no federally certified voting system in the US does that. *Batch-level comparison audits*, which compare electronic tallies for bundles of ballots to hand counts of the votes on those ballots, can be performed by substituting the new test statistic we introduce here into existing batch-level RLA methods.

Section 2 develops RLAs for voting schemes in which each voter may cast at most one vote per party, but possibly several votes in all. Section 3 develops a method applicable when voters may cast several votes among different lists. In both sections, we show how to audit which candidates deserve each party's seats, if a simple plurality system is used for that

[3] We do not address auditing the Belgian allocation of party seats to candidates, nor do we develop RLAs for the complex multi-stage seat allocation in Danish and German parliamentary elections—although our methods could form the basis of such audits. As is, the methods can audit the allocation of most of the seats, but not the "compensatory" rounds.

step, as it is in Danish, Luxembourgish, and Norwegian municipal elections. We illustrate the approach using data from the Danish European Parliamentary Election in 2014. Some countries, including Belgium, use a more complicated algorithm for seating candidates within a party—we do not address that audit.

1.2 "Highest-averages" voting methods

"Highest-averages" methods are party-list proportional representation methods: Each voter chooses a party, and the seats are allocated to parties in proportion to the votes each received. Complications arise from rounding, since seats come in integral numbers. (Complications also arise when voters may cast votes for individual candidates or for more than one party. We address such issues in Section 3.) Throughout the paper, we use "party" and "list" interchangeably.

A list of divisors $d(1), d(2), \ldots, d(S)$ determines a highest-averages method. Starting with the tally $t(p)$ for each party $p = 1, \ldots, P$, seats are allocated by calculating $p_{ps} = t(p)/d(s)$ for $p = 1, \ldots, P$ and $s = 1, \ldots, S$. The S seats go to the parties corresponding to the S largest values of p_{ps}, that is, the *winning set* $\mathcal{W}$ is

$$\mathcal{W} = \{(p, s) : t(p)/d(s) \text{ is one of the } S \text{ largest.}\}$$

Every other candidate loses:

$$\mathcal{L} = \{(p, s) \notin \mathcal{W}\}$$

The number of seats assigned to party p is $\#\{(i, s) \in \mathcal{W} : i = p\}$. Some countries compute all $P \times S$ values of p_{ps}, then choose the largest S entries. Others (such as Luxembourg) derive the same result using iterative calculations, known as Jefferson's or Webster's method.

Notation is summarised in Table 1. Table 1.2 shows how seats were allocated to each coalition in the 2014 Danish EU Parliamentary elections, using D'Hondt. Won seats are shown in bold—these were subsequently distributed among coalition members.

If there are more than S values of p_{ps} greater than or equal to the Sth largest, a tie-breaking rule is used to select S of them. In this case the margin is zero and a full hand count is required. Hence we assume from now on that $\#\mathcal{W} = S$ and $\#\mathcal{L} = S(P - 1)$.

"Highest-averages" methods differ in their choice of divisors. Belgium, Denmark, Luxembourg, and many others use the D'Hondt method, for which $d(i) = i$. Sainte-Laguë, which Germany uses, has divisors $1, 3, 5, 7, \ldots$. Estonia and Norway use variants of D'Hondt and Sainte-Laguë respectively.

2 RLAs for one vote per party

Think of each of the $P \times S$ pairs (p, s) as a *pseudo-candidate* reported to have received p_{ps} votes. The set $\mathcal{W}$ contains the reported winners according to the reported tally. The *reported outcome* is the number of

$$
\begin{aligned}
B &: \text{number of ballots cast in the contest} \\
V &: \text{maximum number of votes per ballot} \\
P &: \text{number of parties} \\
S &: \text{number of seats to be assigned} \\
C_p &: \text{number of candidates in party } p \\
t(p) &: \text{reported total for party } p \\
a(p) &: \text{the actual total for party } p \\
e(p) &\equiv t(p) - a(p), \text{ error for party } p \\
t(p,c) &: \text{reported total for candidate } c \text{ in party } p \\
a(p,c) &: \text{actual total for candidate } c \text{ in party } p \\
e(p,c) &\equiv t(p,c) - a(p,c), \text{ error for candidate } c \text{ in party } p \\
d(s) &: \text{the divisor for column } s \\
p_{ps} &\equiv t(p)/d(s) \\
\pi_{ps} &\equiv a(p)/d(s) \\
\mathcal{W} &: \text{the pairs } (p,s) \text{ with the } S \text{ largest values of } p_{ps} \\
\mathcal{L} &: \text{the pairs } (p,s),\ p = 1,\ldots,P,\ s = 1,\ldots,S \text{ not in } \mathcal{W} \\
\mathcal{W}^P &: \text{the parties } p \text{ that (reportedly) won at least one seat} \\
\mathcal{L}^P &: \text{the parties } p \text{ that (reportedly) lost at least one seat} \\
\mathcal{W}_p &: \text{the candidates } c \text{ in party } p \text{ who were seated} \\
\mathcal{L}_p &: \text{the candidates } c \text{ in party } p \text{ who were not seated}
\end{aligned}
$$

Table 1: Notation

	Count in thousands					
Coalition/party	$t(p)$	/2	/3	/4	/5	/6
A+B+F	**833**	**417**	**278**	**208**	**167**	139
Danish People's	**606**	**303**	**202**	**151**	121	101
C+V	**588**	**294**	**196**	147	118	98
People against EU	**184**	92	61	46	37	31
Liberal Alliance	65	33	22	16	13	11

Table 2: Allocating 13 seats among 5 coalitions using D'Hondt, Danish 2014 EU Parliamentary election.

seats each party gets according to the reported totals $t(p)$, $p = 1, \ldots, P$. The *actual outcome* is the number of seats each party would get according to the actual totals $a(p)$, $p = 1, \ldots, P$. The reported outcome is correct if matches the actual outcome, *i.e.*, if and only if

$$\forall (p_w, s_w) \in \mathcal{W},\ \forall (p_\ell, s_\ell) \in \mathcal{L},\ \pi_{p_w s_w} > \pi_{p_\ell s_\ell}, \tag{1}$$

where $\pi_{ps} \equiv a(p)/d(s)$. Auditing consists of checking those $S^2(P-1)$ inequalities statistically. Some of them are entailed by others because $\pi_{ps} > \pi_{pt}$ for $s < t$ for any method with $d(s) < d(t)$. Hence, for instance, if $\pi_{p_w s_w} > \pi_{p_\ell s_\ell}$, then $\pi_{p_w s_w} > \pi_{p_\ell s}$ for all $s \geq s_\ell$, and $\pi_{p_w s} > \pi_{p_\ell s_\ell}$ for all $s \leq s_w$.

For party p, define

$$s_w(p) \equiv \max\{s : (p, s) \in \mathcal{W}\}$$
$$s_\ell(p) \equiv \min\{s : (p, s) \in \mathcal{L}\}.$$

These are the column indices of the last seat p wins and the first seat p loses, respectively. If p won no seats then $s_w(p)$ doesn't exist; if all p's candidates won then $s_\ell(p)$ doesn't exist. At most S parties can have both winners and losers, so at most $\min(2P, S+P)$ of these exist. Define

$$\mathcal{W}^P \equiv \{p : \exists s \text{ s.t. } (p, s) \in \mathcal{W}\}$$
$$\mathcal{L}^P \equiv \{p : \exists s \text{ s.t. } (p, s) \in \mathcal{L}\}.$$

According to the reported results, these are the parties that won at least one seat and the parties that lost at least one seat, respectively. The inequalities that must be checked by auditing are

$$\forall p \in \mathcal{W}^P,\ \forall q \in \mathcal{L}^P \text{ s.t. } p \neq q,\ \pi_{p, s_w(p)} > \pi_{q, s_\ell(q)}. \tag{2}$$

2.1 Ballot-polling Audits

Assumption We assume in this section that the voting rules allow voters to cast at most one vote for at most one party. (A risk-limiting ballot-polling method when voters may cast votes for more than one party or more than one vote per party is given below in Section 3.2.1.)

We will modify the ballot-polling audit method introduced by Lindeman, Stark, and Yates [2012]. Consider a pair of pseudo-candidates $(p_w, s_w) \in \mathcal{W}$ and $(p_\ell, s_\ell) \in \mathcal{L}$, with $p_w \neq p_\ell$. We want to use a random sample to collect and assess evidence regarding whether $\pi_{p_w s_w} > \pi_{p_\ell s_\ell}$. That inequality amounts to $a(p_w)/d(s_w) > a(p_\ell)/d(s_\ell)$, *i.e.*,

$$a(p_w) > a(p_\ell) \frac{d(s_w)}{d(s_\ell)}. \tag{3}$$

Suppose inequality (3) holds. Imagine drawing ballots at random. Let A_p be the event that a randomly selected ballot shows a vote for party p. Then $\Pr(A_p) = a(p)/B$. If the outcome is correct (and if at least one ballot was cast for party p_w or for p_ℓ),

$$\Pr(A_{p_w}) \geq \frac{d(s_w)}{d(s_\ell)} \Pr(A_{p_\ell}),$$

which implies that

$$\Pr(A_{p_w}|A_{p_w} \cup A_{p_\ell}) \geq \frac{d(s_w)}{d(s_\ell)} \Pr(A_{p_\ell}|A_{p_w} \cup A_{p_\ell}),$$

since $A_{p_w} \subset A_{p_w} \cup A_{p_\ell}$ and $A_{p_\ell} \subset A_{p_w} \cup A_{p_\ell}$. That is, the conditional probability $\pi_{p_w|p_w p_\ell}$ that a randomly selected ballot shows a vote for party p_w given that it shows a vote either for p_w or p_ℓ must be at least $d(s_w)/d(s_\ell)$ times the conditional probability $\pi_{p_\ell|p_w p_\ell}$ that such a ballot shows a vote for party p_ℓ. Those two conditional probabilities sum to 100%. Hence, for the outcome to be correct, we need

$$\begin{aligned} \pi_{p_w|p_w p_\ell} &> (1 - \pi_{p_w|p_w p_\ell}) d(s_w)/d(s_\ell) \\ \pi_{p_w|p_w p_\ell}(1 + d(s_w)/d(s_\ell)) &> d(s_w)/d(s_\ell) \\ i.e., \quad \pi_{p_w|p_w p_\ell} &> \frac{d(s_w)}{d(s_\ell) + d(s_w)}, \end{aligned} \tag{4}$$

and

$$\pi_{p_\ell|p_w p_\ell} < 1 - \frac{d(s_w)}{d(s_\ell) + d(s_w)}. \tag{5}$$

Now,

$$\pi_{p_w|p_w p_\ell} \equiv \frac{a(p_w)}{a(p_w) + a(p_\ell)}$$

and

$$\frac{t(p_w)}{t(p_w) + t(p_\ell)} > \frac{d(s_w)}{d(s_\ell) + d(s_w)}.$$

We can use Wald's sequential probability ratio test [Wald, 1945] to test the null hypothesis that

$$\frac{a(p_w)}{a(p_w) + a(p_\ell)} \leq \frac{d(s_w)}{d(s_\ell) + d(s_w)}$$

against the alternative hypothesis that

$$\frac{a(p_w)}{a(p_w) + a(p_\ell)} \geq \frac{t(p_w)}{t(p_w) + t(p_\ell)}.$$

To reject the null hypothesis is to confirm that $\pi_{p_w s_w} > \pi_{p_\ell s_\ell}$. In a single draw from the population of ballots, conditional on the event that the ballot shows a vote for either p_w or p_ℓ, the likelihood ratio for the alternative to the null is

$$\frac{\frac{t(p_w)}{t(p_w)+t(p_\ell)}}{\frac{d(s_w(p_w))}{d(s_w(p_w))+d(s_\ell(p_\ell))}}$$

if the ballot shows a vote for p_w. Under the same condition, the likelihood ratio for the alternative to the null is

$$\frac{1 - \frac{t(p_w)}{t(p_w)+t(p_\ell)}}{1 - \frac{d(s_w(p_w))}{d(s_w(p_w))+d(s_\ell(p_\ell))}}$$

if the ballot shows a vote for p_ℓ. Using this likelihood ratio with Wald's sequential probability ratio test [Wald, 1945] gives the following algorithm for an RLA with risk limit α:

1. Select the risk limit $\alpha \in (0, 1)$, and M, the maximum number of ballots to audit before proceeding to a full hand count. Define

$$\gamma^{+}_{ps_w(p)qs_\ell(q)} \equiv \frac{t(p)}{t(p)+t(q)} \cdot \frac{d(s_w(p)) + d(s_\ell(q))}{d(s_w(p))}$$

and

$$\gamma^{-}_{ps_w(p)qs_\ell(q)} \equiv \left(1 - \frac{t(p)}{t(p)+t(q)}\right) \times \left(1 - \frac{d(s_w(p)) + d(s_\ell(q))}{d(s_w(p))}\right).$$

Set $T_{ps_w(p)qs_\ell(q)} = 1$ for all $p \in \mathcal{W}^P$ and $q \in \mathcal{L}^P$, $p \neq q$. Set $m = 0$.

2. Draw a ballot uniformly at random with replacement from those cast in the contest and increment m.
3. If the ballot shows a valid vote for a reported winner $p \in \mathcal{W}^P$, then for each $q \neq p$ in $\mathcal{L}^P$ that did not receive a valid vote on that ballot multiply $T_{ps_w(p)qs_\ell(q)}$ by $\gamma^{+}_{ps_w(p)qs_\ell(q)}$. Repeat for all such p.
4. If the ballot shows a valid vote for a reported loser $q \in \mathcal{L}^P$, then for each $p \neq q$ in $\mathcal{W}^P$ that did not receive a valid vote on that ballot, multiply $T_{ps_w(p)qs_\ell(q)}$ by $\gamma^{-}_{ps_w(p)qs_\ell(q)}$. Repeat for all such q.
5. If any $T_{ps_w(p)qs_\ell(q)} \geq 1/\alpha$, reject the corresponding null hypothesis for each such $T_{ps_w(p)qs_\ell(q)}$. Once a null hypothesis is rejected, do not update its $T_{ps_w(p)qs_\ell(q)}$ after subsequent draws.
6. If all null hypotheses have been rejected, stop the audit: The reported results stand. Otherwise, if $m < M$, return to step 2.
7. Perform a full hand count; the results of the hand count replace the reported results.

Because

$$\frac{t(p_w)}{t(p_w)+t(p_\ell)} > \frac{d(s_w)}{d(s_\ell)+d(s_w)},$$

$Tp_w s_w(p_q)p_\ell s_\ell(p_\ell)$ increases when a ballot with a vote for p_w is drawn and decreases when a ballot for p_ℓ is drawn. If all the alternative hypotheses are true, the values of all the T will tend to increase. If any of the null hypotheses is true, the chance is less than α that the corresponding value of T will ever exceed $1/\alpha$. Hence, as discussed in Lindeman et al. [2012], if any of the null hypotheses is true, despite the fact that we are comparing many pairs of probabilities, there is a large chance that the procedure will require a full hand count: The issue of multiplicity does not arise.

2.2 Comparison audits

Assumption We continue to assume that the voting rules allow voters to cast at most one vote per party, but now we allow votes for multiple parties. (This assumption is relaxed in Section 3.2.2.)

Our approach is similar to the maximum (in-contest) relative overstatement of pairwise margins introduced by Stark [2008b], but with weights in the numerator to account for the fact that a vote for party p amounts to (differing) fractional votes for all the pseudo-candidates in row p.

First, we will transform the problem slightly so that we can use MICRO. We seek a simple sufficient condition for the correctness of the outcome in terms of $e(p)$, $p = 1, \ldots, P$; that is, a condition on the errors in the reported tally that ensures $\pi_{p_w s_w} > \pi_{p_\ell s_\ell}$, $\forall (p_w, s_w) \in \mathcal{W}$ and $\forall (p_\ell, s_\ell) \in \mathcal{L}$.

Suppose there is some $p_\ell \in \mathcal{L}$ and $p_w \in \mathcal{W}$ for which $\pi_{p_w s_w} \leq \pi_{p_\ell s_\ell}$; that is, some seat has been misallocated. Then

$$\begin{aligned}
\pi_{p_\ell s_\ell} - \pi_{p_w s_w} &\geq 0 \\
\pi_{p_\ell s_\ell} - p_{p_\ell s_\ell} - (\pi_{p_w s_w} - p_{p_w s_w}) &\geq p_{p_w s_w} - p_{p_\ell s_\ell} \\
\frac{(p_{p_w s_w} - \pi_{p_w s_w}) - (p_{p_\ell s_\ell} - \pi_{p_\ell s_\ell})}{p_{p_w s_w} - p_{p_\ell s_\ell}} &\geq 1.
\end{aligned}$$

A little algebra using the definition $p_{ps} \equiv t(p)/d(s)$ shows that the outcome must therefore be correct if

$$\text{MICRO} \equiv \max_{(p_w, s_w) \in \mathcal{W},\, (p_\ell, s_\ell) \in \mathcal{L}} \frac{d(s_\ell)e(p_w) - d(s_w)e(p_\ell)}{d(s_\ell)t(p_w) - d(s_w)t(p_\ell)} < 1.$$

It suffices to take the maximum over $p_w \neq p_\ell$: a party cannot lose a seat to itself.

Let $e_b(p)$ denote the error in the tally of the vote for party p on ballot b. Then $e(p) = \sum_{b=1}^{B} e_b(p)$. Since the sum of maxima dominates the maximum of sums, MICRO < 1 if

$$\sum_{b=1}^{B} \max_{(p_w, s_w) \in \mathcal{W},\, (p_\ell, s_\ell) \in \mathcal{L}: p_w \neq p_\ell} \frac{d(s_\ell)e_b(p_w) - d(s_w)e_b(p_\ell)}{d(s_\ell)t(p_w) - d(s_w)t(p_\ell)} < 1. \tag{6}$$

We now derive a test of hypothesis that MICRO ≥ 1 based on the Kaplan-Wald approach, derived in Appendix A. The test can be modified to use reported results for bundles of ballots rather than individual ballots, at the expense of some bookkeeping; we do not present that generalization here, because for typical bundle sizes and modest margins, it offers little or no advantage over ballot-polling audits, which have far lower set-up costs.

Although one ballot may have been miscounted in a way that affects more than two parties, we need only count the errors that have the largest combined effect on the margin between two pseudo-candidates, because we are summing the maximum effect in the test. Since $|e_b(p)| \leq 1$, the largest possible contribution of any ballot to the left hand side of (6) is

$$u \equiv \max_{w \in \mathcal{W}^P,\, \ell \in \mathcal{L}^P: w \neq \ell} \frac{d(s_\ell(\ell)) + d(s_w(w))}{d(s_\ell(\ell))t(p_w(w)) - d(s_w(w))t(p_\ell(\ell))}. \tag{7}$$

The Kaplan-Wald method requires sampling ballots independently with a probability of selecting each ballot proportional to an upper bound on MICRO for that ballot. Using u as the upper bound on MICRO for *every* ballot results in sampling ballots with equal probabilities—and is conservative.

The following algorithm gives RLA at risk limit α. We assume as before that a compliance audit has shown the audit trail to be sufficiently complete and accurate that a full hand count would show the correct electoral outcome.

The constant γ is a tuning parameter that trades off effort when the cast vote records are error-free against the effort when the cast vote records have errors. The larger γ is (within $[0,1]$), the smaller the sample will need to be to confirm the outcome when none of the cast vote records is discovered to have error, but the larger the sample will need to be if the audit uncovers errors.

1. Select the risk limit $\alpha \in (0,1)$; M, the maximum number of ballots to audit before proceeding to a full hand count; and $\gamma \in (0,1)$. Calculate u and $U = Bu$, the maximum total overstatement. Set $m = 0$.
2. Draw a ballot uniformly at random with replacement from those cast in the contest and increment m.
3. Find MICRO for the selected ballot and divide it by u. Denote the quotient D_m.
4. Calculate $\beta = \prod_{i=1}^{m} \left[\gamma \frac{1-D_i}{1-1/U} + 1 - \gamma\right]$.
5. If $\beta > 1/\alpha$, stop the audit: The outcome is confirmed at risk limit α.
6. If $m < M$, return to step 2.
7. Perform a full hand count; the results of the hand count replace the reported results.

It is a theorem that if any seat was misallocated, the chance this algorithm proceeds to a full hand count is at least $1 - \alpha$: the risk limit is α. Smaller values of γ reduce the increase in workload when discrepancies are found, but increase the workload when no discrepancies are found. The risk limit is conservative regardless.

The method can be simplified and still remain conservative if we replace step 3 by

3') If the selected ballot agrees perfectly with the cast vote record, set $D_m = 0$; otherwise, set $D_m = 1$.

That substitution eliminates the need for any algebra when a discrepancy is discovered, and makes the calculation in step 4 simple. However, it can require inspecting far more ballots when the outcome is correct and discrepancies are observed, because each discrepancy results in multiplying β by $1 - \gamma$.

2.3 Applicability

This method could be used immediately for auditing the number of seats obtained by each list wherever voters may cast only one vote for a list, for example in Danish municipal elections and in Belgium. (See the next section for auditing the candidates assigned to each seat.)

It could also be used in Danish and German parliamentary elections to audit the number of seats obtained by each list in the first (pure) round of D'Hondt tallying. In both countries, the technique would have to be augmented to deal with their complex processes for allocating "compensatory" seats in addition to the D'Hondt count.

Many countries also impose a threshold for parliamentary representation. Some (including Estonia) allow candidates or parties who have exceeded a threshold to be seated immediately, before the D'Hondt count. These could be checked in a straightforward simultaneous audit.

2.4 Illustration: 2014 EU Parliamentary Election in Denmark

Reported results for the 2014 EU Parliamentary election in Denmark are in table 1.2.

An IPython notebook with the data and algorithms is in appendix B and available at XXX. For $\gamma = 0.95$, the allocations of seats to coalitions could have been confirmed at 99.9% confidence ($\alpha = 0.001$ risk limit) by inspecting 1903 ballots—if the audit did not find any errors in that sample.

3 Extension to individual-candidate variants

Many countries allow individual candidate votes. Details vary, but in broad brush, instead of or in addition to choosing a party, voters may select or delete individual candidates. The allocation of seats to parties is as above, based on a combination of party list votes and individual candidate votes. The individual candidate votes are used to decide which candidates in the party are seated.

The electoral outcome can be wrong—the wrong individuals can get seats—either because the parties get the wrong number of seats or because the t candidates within a party that was correctly allocated t seats are not the correct candidates to seat. In many countries, the t candidates in a party who are seated are the t who received the most votes. In that case, we need to test whether every party got the right number of seats and whether, for each party that received at least one seat, the t candidates who reportedly received the most votes really did receive the most votes. The latter amounts to auditing a collection of plurality contests with multiple winners [Stark, 2009]. Below, we extend ballot-polling audits to cover this case.

3.1 Single-list votes plus candidates

In parliamentary elections in Denmark, Belgium, Germany, Estonia, and Norway, voters cast a single party-list vote and may also vote for individual candidate(s) within that list.[4] In these cases, the audit of the seats

[4]This idea is expressed slightly differently in each country. In Denmark, a voter selects either a candidate or a party list. A vote for a candidate is equivalent to a vote for their party list for the purposes of the D'Hondt allocation, but also counts towards that candidate's individual tally for the purposes of assigning seats to candidates within a party. In Germany, voters may select an individual candidate directly (which does not influence the allocation of seats to parties in the Sainte-Laguë count) and then also cast a party-list vote. In Belgium,

allocated per party by the highest-averages count is exactly the same as in the pure case. The audit of which candidates should be seated within each party can be performed simultaneously using the same random sample of ballots by combining tests of pairwise majorities within each party with the test of weighted majorities across parties:

- For ballot-polling, this only requires including additional test statistics for each (seated, non-seated) pair within a party, following Lindeman et al. [2012]: for each pair, we seek strong evidence that the seated candidate received more than half of the votes on ballots that contain votes for either or both candidates.
- For ballot-level comparison audits, we can combine the tests that the seated candidates each received more votes than any of the non-seated candidates by using the maximum across-contest relative overstatement (MACRO) across the pairwise within-party contests, exactly as described by [Stark, 2009]. The Kaplan-Wald method can be used to test the hypothesis that MACRO ≥ 1.

These approaches solve the auditing problem for Denmark and Germany,[5] but not Belgium, which would require a specialized technique tailored to its complicated allocation algorithm.

3.2 Multiple list votes

Assumption This section considers rules that allow a voter to cast multiple votes per party or votes for more than one party. For instance, some countries allow voters to endorse several candidates, who need not be in the same party. For the purposes of a highest-averages method, this is equivalent to giving each voter several votes, which she may distribute among several lists. The highest-averages count then proceeds exactly as in the pure case, except there may be several votes per voter.

The ballot-level comparison RLA of section 2.2 can be modified easily to allow for this possibility—see Section 3.2.2. However, the basic ballot-polling RLA of section 2.1 cannot, even though a comparably simple ballot-polling method works in plurality contests where voters may cast votes for more than one candidate.[6] We develop a different method below in section 3.2.1

Rules vary widely among such systems. For instance, in Luxembourg, voters may choose either a party vote or a candidate vote. In the latter, they may cast up to S votes in total, including up to 2 votes for any single

voters may choose either a party list or an arbitrary number of candidates from the same list, which again is equivalent for the purposes of the D'Hondt tally. Then candidates are seated within parties using a complicated algorithm that combines the voters' and the parties' choices.

[5]This applies only to the first-round of Sainte-Laguë, not the second-round that allocates extra seats in the Bundestag using a different system.

[6]In plurality contests, the basic ballot polling audit checks whether, among ballots that list exactly one of two candidates, one candidate has the majority. Ballots that show both candidates can be ignored. But when a voter can cast more than one vote per party, Party p can have $d(q)/d(p)$ times as many votes as party q among ballots that list exactly one of the two parties, but still not have $d(q)/d(p)$ times as many votes in all, so that conditioning does not yield a valid test.

candidate. Party votes are interpreted as one vote for every candidate on the party list. Party totals are used to allocate seats to parties by D'Hondt; individual votes are used to allocate seats within each party.

Auditing the allocation of seats to candidates requires a method appropriate for the tallying scheme. In Luxembourg this is simple (multi-winner) plurality; in Norwegian municipal elections it is a plurality variant weighted by party selections. Both ballot-polling and ballot-level comparison RLAs can be extended to audit simultaneously how many seats each party gets and which candidates get each party's seats, assuming the latter done by simple plurality. For illustration, we present ballot-polling and comparison audits for the Luxembourgish system.

3.2.1 Ballot-polling audit

Developing a RLA for the Luxembourgish system requires a different approach than that of section 2.1. Voters may cast up to S votes, and up to 2 per candidate, so the probability that a randomly selected ballot shows a vote for a given party or candidate is not proportional to the number of votes for that party or candidate.

This method uses differences in expected values of the number of votes for different parties (normalized by the appropriate column divisors $d(\cdot)$) or for different candidates. We treat a party-list vote as a set of individualised votes for all candidates in that party. Suppose we select a ballot uniformly at random from the B ballots cast. Let $V_{p,c}$ denote the number of votes for candidate c in party p on that ballot and let $V_p \equiv \sum_{c=1}^{C_p} V_{p,c}$ denote the total number of votes for party p on that ballot. Then the expected value of $V_{p,c}$ is

$$\mathbb{E}V_{p,c} = t(p,c)/B \text{ and } \mathbb{E}V_p = t(p)/B.$$

Moreover,

$$\mathbb{E}(V_p/d(s) - V_q/d(t)) = \frac{t(p)/d(s) - t(q)/d(t)}{B}. \tag{8}$$

The allocation of seats to parties is therefore correct if

$$\forall p \in \mathcal{W}^P,\ \forall q \in \mathcal{L}^P \text{ s.t. } p \neq q,\ \mathbb{E}\left(\frac{V_p}{d(s_w(p))} - \frac{V_q}{d(s_\ell(q))}\right) > 0. \tag{9}$$

The allocation of seats to candidates in those parties is also correct if

$$\forall p \in \mathcal{W}^P,\ c_w \in \mathcal{W}_p,\ c_\ell \in \mathcal{L}_p,\ \mathbb{E}(V_{p,c_w} - V_{p,c_\ell}) > 0. \tag{10}$$

Voting rules for a particular country impose constraints that imply lower and upper bounds on the combinations of random variables on the left-hand sides of (9) and (10). Let X_i denote any of those left-hand sides, calculated for the ith draw. (Draws are random, independent, and uniformly distributed.) Let x_+ and x_- denote the upper and lower bounds respectively. For example, in the Luxembourgish system, the rules require $V_{p,c} \leq 2$ and $V_p \leq S$, so if X_i denotes

$$\frac{V_p}{d(s_w(p))} - \frac{V_q}{d(s_\ell(q))}$$

for the ith draw (Eq. (9)), then x_+ is $S/d(s_w(p))$ and x_- is $-S/d(s_\ell(q))$.

We know a priori that $x_- \leq X_i \leq x_+$; we wish to test the hypothesis that $\mathbb{E}X_i \leq 0$. Rejecting that hypothesis for all the left-hand sides confirms the seat allocation. Let $\tilde{X}_i \equiv (X_i + x_-)/(x_+ - x_-)$. Then $\tilde{X}_i \in [0,1]$, and the condition $\mathbb{E}X_i \leq 0$ is equivalent to the condition $\mathbb{E}\tilde{X}_i \leq t \equiv x_-/(x_+ - x_-)$. Imagine drawing n ballots, resulting in $\{\tilde{X}_i\}_{i=1}^n$ independent and identically distributed on $[0,1]$. Define

$$\mathrm{LR} \equiv \prod_{i=1}^{n} \left[\gamma \frac{\tilde{X}_i}{t} + 1 - \gamma \right]. \tag{11}$$

Much the same proof as in appendix A[7] shows that if $\mathbb{E}\tilde{X}_i \leq t$, $\Pr\{\mathrm{LR} > 1/\alpha\} \leq \alpha$ for any n. We can use this result to test all the conditions (9) and (10) with a single sample. Multiplicity is not a concern because the audit proceeds to a full hand count if *any* null hypothesis is not rejected.

3.2.2 Ballot-level comparison audit

The algorithm of section 2.2 can audit the number of seats allocated to parties in the case of allowing up to V votes per party per voter, except that the upper bound u_m on the maximum possible value of MICRO for a single ballot is V times as large as that in Equation 7. For Luxembourg, the maximum votes per ballot is the number of available seats, so $u_m = Su$. To include the competition for seats among members of the same party, we need only consider that competition to be a collection of pairwise elections between all candidates in a party who were awarded seats and all who were not. Table 1 outlines the notation.

The definition of MACRO incorporating both kinds of error is:

$$\mathrm{MACRO}_{\mathrm{multi}} \equiv \max \left\{ \mathrm{MICRO}, \max_{p \in \mathcal{W}^P, c_w \in \mathcal{W}_p, c_\ell \in \mathcal{L}_p} \frac{e(p, c_w) - e(p, c_\ell)}{t(p, c_w) - t(p, c_\ell)} \right\}.$$

If $\mathrm{MACRO}_{\mathrm{multi}} < 1$, the allocation of seats to parties and the allocation of seats to candidates within parties are all correct.

3.2.3 Logistical and statistical concerns

If relatively few votes separate a seated candidate from a candidate in the same party who is not seated, the sample sizes needed to attain reasonable risk limits using the methods presented above will be very large. If it is possible to divide the ballots into (overlapping) subsets that contain only the ballots cast for a particular party, and to sample directly from those subsets, it may be possible to reduce sample sizes, depending on the margins compared to the number of ballots in each subset. Auditing the allocation of seats within parties separately from auditing the allocation of seats to parties also raises issues of multiple testing, which will tend to increase the required sample size to attain a given risk limit.

[7] The proof appears in sketch form in `http://printmacroj.com/martMean.htm`, last accessed 10 November 2013.

3.3 Audit Summary

We have presented ballot-polling and ballot-level comparison RLAs for all highest-averages proportional representation methods, including those in which voters select a single party list and those in which they may cast some votes for each of several parties and more than one vote for the same party or candidate. We have shown that the same sample can be used to check that the right candidates were seated within each party, at least for (the many) countries that use plurality or a simple variant to allocate seats to candidates. The methods need modifications to check the "compensatory" rounds in German and Danish parliamentary elections, which do not use a highest-averages method, and for auditing which candidates get the party's seats in non-plurality systems such as Belgian and Norwegian parliamentary elections.

4 Conclusion

Highest-averages methods include many party-list proportional representation methods, implemented differently in different countries—and sometimes in different ways in a single country. The pure versions of these methods are amenable both to efficient risk-limiting audits and to complete homomorphic tallying. We develop methods for several variants, some of which are particularly important because the country uses or plans to use electronic voting. In particular, we illustrate risk-limiting audits for Denmark and privacy-preserving universally verifiable tallying for Norway. The methods allow election outcomes of D'Hondt, Sainte-Laguë, and variants to be verified.

5 Acknowledgements

Many thanks to Christian Bull, Sven Heiberg, Carsten Schürmann, Kristian Gjøsteen, Olivier Pereira, Melanie Volkamer, Peter Ryan and Rui Joaquim.

References

Bevoting study of electronic voting systems, 2007. Available at `http://www.ibz.rrn.fgov.be/fileadmin/user_upload/Elections/fr/presentation/bevoting-2_gb.pdf`.

J. Benaloh, D. Jones, E. Lazarus, M. Lindeman, and P.B. Stark. Soba: Secrecy-preserving observable ballot-level audit. In *proc. Proc. USENIX Accurate Electronic Voting Technology Workshop*, 2011.

M. Gallagher. Comparing proportional representation electoral systems: Quotas, thresholds, paradoxes and majorities. *British Journal of Political Science*, 22(4):469–496, 1992.

J.L. Hall, L.W. Miratrix, P.B. Stark, M. Briones, E. Ginnold, F. Oakley, M. Peaden, G. Pellerin, T. Stanionis, and T. Webber. Implementing risk-limiting post-election audits in California. In *Proc. 2009 Electronic Voting Technology Workshop/Workshop on Trustworthy Elections (EVT/WOTE '09)*, Montreal, Canada, August 2009. USENIX. URL `http://www.usenix.org/event/evtwote09/tech/full_papers/hall.pdf`.

H.M. Kaplan. A method of one-sided nonparametric inference for the mean of a nonnegative population. *The American Statistician*, 41:157–158, 1987.

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

M. Lindeman, P.B. Stark, and V. Yates. BRAVO: Ballot-polling risk-limiting audits to verify outcomes. In *Proceedings of the 2011 Electronic Voting Technology Workshop / Workshop on Trustworthy Elections (EVT/WOTE '11)*. USENIX, 2012.

Panel on Nonstandard Mixtures of Distributions. *Statistical models and analysis in auditing: A study of statistical models and methods for analyzing nonstandard mixtures of distributions in auditing*. National Academy Press, Washington, D.C., 1988.

A.D. Sarwate, S. Checkoway, and H. Shacham. Risk-limiting audits and the margin of victory in nonplurality elections. *Politics, and Policy*, 3 (3):29–64, 2013.

P.B. Stark. Conservative statistical post-election audits. *Annals of Applied Statistics*, 2008a.

P.B. Stark. A sharper discrepancy measure for post-election audits. *Ann. Appl. Stat*, 2:982–985, 2008b.

P.B. Stark. Efficient post-election audits of multiple contests: 2009 California tests. `http://ssrn.com/abstract=1443314`, 2009. 2009 Conference on Empirical Legal Studies.

P.B. Stark and D.A. Wagner. Evidence-based elections. *IEEE Security and Privacy*, 10:33–41, 2012.

A. Wald. Sequential tests of statistical hypotheses. *Ann. Math. Stat.*, 16: 117–186, 1945.

A. Wald. *Sequential Analysis*. Dover Publications, Mineola, New York, 2004.

A The Kaplan-Wald Method

This section combines ideas from dollar-unit sampling as used in financial auditing [Panel on Nonstandard Mixtures of Distributions, 1988] with a technique described in H.M. Kaplan's website, `http://printmacroj.com/martMean.htm`.[8] Kaplan's work fleshes out an idea due to Wald [Wald, 1945, 2004], and is closely related to a technique presented in Kaplan [1987]. We have a population of N items. Item j has a value x_j between 0 and a known upper bound $u_j > 0$. We wish to estimate the population total $T = \sum_{j=1}^{N} x_j$.

Define $d_j \equiv x_j/u_j$, for $j = 1, \ldots, N$. Each d_j is necessarily between 0 and 1. Let $U = \sum_{j=1}^{N} u_j$. We will make n independent random draws with replacement from the population; the probability of selecting item j is $p_j \equiv u_j/U$ in each draw. That is, the chance of selecting item j is proportional to its upper bound.

Let $J(i)$ be the index of the item selected on the ith draw. Let D_i be $d_{J(i)}$, the value of d for the item selected on the ith draw. For instance, if the second draw gives the fifth item, then $J(2) = 5$ and $D_2 = d_{J(2)} = d_5$. The chance that $J(i) = j$ is p_j. The expected value of D_i is

$$\begin{aligned}
\mathbb{E}D_i &= \sum_{j=1}^{N} d_j \Pr(J(i) = j) \\
&= \sum_{j=1}^{N} d_j (u_j/U) \\
&= \sum_{j=1}^{N} (x_j/u_j) \times (u_j/U) \\
&= \sum_{j=1}^{N} x_j/U \\
&= T/U.
\end{aligned}$$

Hence,

$$\mathbb{E}\frac{1}{n}\sum_{i=1}^{n} D_i = \frac{1}{n} nT/U = T/U. \tag{12}$$

That is, the average of the n draws $\{D_i\}_{i=1}^n$ is an unbiased estimator of the population total T as a fraction of the total upper bound U. Equivalently, U times the average of $\{D_i\}_{i=1}^n$ is an unbiased estimate of the population total T.

The expected value of D_i is T/U, which is unknown since T is unknown. For the purpose of conducting a risk-limiting audit, we want to test the hypothesis that $T \geq 1$ (equivalently, that $T/U \geq 1/U$). We will derive a method based on Wald's sequential probability ratio test [Wald, 1945, 2004], following an idea of Harold Kaplan, based in turn on a remark in Wald [2004].[9] Note that if $U < 1$, we do not need to audit: the

[8] Last accessed 10 November 2013.

[9] See `http://printmacroj.com/martMean.htm`. Last accessed 10 November 2013.

maximum possible value of MICRO is less than 1, so the outcome must be correct. We therefore assume that $U \geq 1$.

The likelihood ratio of the simple hypothesis H_1 to the simple hypothesis H_0 is the probability of observing the data that actually were observed on the assumption that H_1 is true, divided by the probability of observing the data that were actually observed on the assumption that H_0 is true:

$$\text{likelihood ratio} \equiv \frac{\Pr(\text{observed data if } H_1 \text{ is true})}{\Pr(\text{observed data if } H_0 \text{ is true})}. \tag{13}$$

The probability of observing the data actually observed will tend to be higher for whichever hypothesis is in fact true, so the likelihood ratio will tend to be greater than 1 if H_1 is true, and will tend to be less than 1 if H_0 is true. The more observations we make, the more probable it is that the resulting likelihood ratio will be small if H_0 is true. Wald [1945] showed that if H_0 is true, then the probability is at most α that the likelihood ratio is ever greater than $1/\alpha$, no matter how many observations are made.

Let $\tilde{D}_i \equiv 1 - D_i$, and let $\tilde{d}_j \equiv 1 - d_j$. Then the probability distribution of $\tilde{D}_i$ is

$$f(d) \equiv \sum_{j=1}^{N} p_j \delta(d - \tilde{d}_j)),$$

where $\delta(\cdot)$ is the Dirac delta function. Under the hypothesis that $T = t$, the expected value of $\tilde{D}_i$ is $1 - t/U$, so the expected value of $(1 - t/U)^{-1}\tilde{D}_i$ is 1. That is,

$$\int_{d=0}^{1} (1 - t/U)^{-1} d f(d) = \sum_{j=1}^{N} (1 - t/U)^{-1} \tilde{d}_j p_j = 1. \tag{14}$$

Let $\gamma \in [0, 1]$ be a fixed number. Because $\sum_j p_j = 1$, it follows that if $T = t$,

$$\begin{aligned}
&\mathbb{E}(\gamma(1 - t/U)^{-1}\tilde{D}_i + (1 - \gamma)) \\
&= \frac{\gamma}{1 - t/U} \mathbb{E}\tilde{D}_i + (1 - \gamma) \\
&= \frac{\gamma}{1 - t/U} (1 - t/U) + 1 - \gamma \\
&= \gamma \cdot 1 + (1 - \gamma) \\
&= 1.
\end{aligned}$$

Now,

$$\mathbb{E}(\gamma(1 - t/U)^{-1}\tilde{D}_i + (1 - \gamma)) \equiv \sum_{j=1}^{N} (\gamma(1 - t/U)^{-1}\tilde{d}_j + 1 - \gamma) p_j. \tag{15}$$

Let

$$g_{j,t,\gamma} \equiv (\gamma(1 - t/U)^{-1}\tilde{d}_j + 1 - \gamma) p_j, \quad j = 1, \ldots, N. \tag{16}$$

Since $t/U \in [0,1]$ and all $\{\tilde{d}_j\}$ are nonnegative, it follows from (15) and (**??**) that $g_{j,t,\gamma} \geq 0$ and

$$\sum_{j=1}^{N} g_{j,t,\gamma} = 1. \tag{17}$$

That is, $\sum_{j=1}^{N} g_{j,t,\gamma}\delta(d - \tilde{d}_j)$ is a probability distribution. Let F be a random variable with $\Pr\{F = \tilde{d}_j\} = g_{j,t,\gamma}$. Since $\mathbb{E}\tilde{D}_i = 1 - t/U \geq 0$,

$$\begin{aligned}
\mathbb{E}F &= \sum_{j=1}^{N} \Big(\gamma(1-t/U)^{-1}\tilde{d}_j + 1 - \gamma\Big)\tilde{d}_j p_j \\
&= \frac{\gamma}{1-t/U}\sum_{j=1}^{N}\tilde{d}_j^2 p_j + (1-\gamma)\mathbb{E}\tilde{D}_i \\
&= \frac{\gamma}{1-t/U}\mathbb{E}\tilde{D}_i^2 + (1-\gamma)\mathbb{E}\tilde{D}_i \\
&\geq \frac{\gamma}{1-t/U}(\mathbb{E}\tilde{D}_j)^2 + (1-\gamma)\mathbb{E}\tilde{D}_i \\
&= \gamma\mathbb{E}\tilde{D}_i + (1-\gamma)\mathbb{E}\tilde{D}_i = \mathbb{E}\tilde{D}_i,
\end{aligned}$$

where the penultimate step follows from Jensen's inequality.

If the data allow us to reject the hypothesis H_0 that $\{\tilde{D}_i\}$ all have the same probability mass function f (for which $\mathbb{E}\tilde{D}_i = 1 - t/U$) in favor of the alternative hypothesis H_1 that $\{\tilde{D}_i\}$ all have the probability mass function $g_{t,\gamma}$ (for which $\mathbb{E}\tilde{D}_i > 1 - t/U$), we have strong statistical evidence that $\mathbb{E}D_i < t/U$. Since $\mathbb{E}D_i < t/U$ is a sufficient condition for the electoral outcome to be correct, rejecting H_0 means the audit can stop: The data gave strong evidence that the election outcome is correct.

Recall that $J(i)$ is the index of the item selected on the ith draw. For n independent observations $\{\tilde{D}_i\}_{i=1}^n$, the likelihood ratio of H_1 to H_0 is

$$\begin{aligned}
\mathrm{LR} &= \frac{\Pr(\text{observed data if } H_1 \text{ is true})}{\Pr(\text{observed data if } H_0 \text{ is true})} \\
&= \frac{\prod_{i=1}^{n}\Big[\gamma(1-t/U)^{-1}\tilde{D}_i + 1 - \gamma\Big]p_{J(i)}}{\prod_{i=1}^{n} p_{J(i)}} \\
&= \prod_{i=1}^{n}\left[\gamma\frac{1-D_i}{1-t/U} + 1 - \gamma\right].
\end{aligned} \tag{18}$$

The dependence on $\{p_j\}$ in the numerator and denominator cancel fortuitously: The validity of the test does not depend on any assumptions about the population $\{d_j\}$ of values. Equation (18) motivates the introduction of γ: For $\gamma = 1$, the likelihood ratio would forever be 0 if even a single observed value of D_i were equal to 1.

To conduct a risk-limiting audit, we take $t = 1$ in (18). If in fact $T \geq 1$, Wald's sequential probability ratio test establishes that the chance that the likelihood ratio is ever larger than $1/\alpha$ is at most α, no matter what the population $\{d_j\}$ of values may be. If we continue to inspect ballots until $\mathrm{LR} > 1/\alpha$—or until we have inspected all the ballots—the chance the audit will stop short of a full hand count if the outcome is wrong is less than α.

USENIX Journal of Election Technology and Systems (JETS)
Volume 3, Number 1 • December 2014
www.usenix.org/jets/issues/0301

B Python code for the European Union Parliamentary election in Denmark

```
import math
import numpy as np
import scipy
from scipy.stats import binom
import pandas as pd
#
def dHondt(partyTotals, seats, divisors):
    '''
    allocate <seats> seats to parties according to <partyTotals> votes,
    using D'Hondt proportional allocation with <weights> divisors

    Input:
        partyTotals: list of total votes by party
        seats:       total number of seats to allocate
        divisors:    divisors for proportional allocation.
                     For d'Hondt, divisors are 1, 2, 3, ...

    Returns:
        partySeats:  list of number of seats for each party
        seated:      list of tuples--parties with at least one seat,
                     number of votes that party got,
                     and divisor for last seated in the party
        notSeated:   list of tuples--parties with at least one lost seat,
                     number of votes that party got,
                     and divisor for the first non-seated in the party
        pseudoCandidates: matrix of votes for each pseudocandidate
    '''
    pseudoCandidates = np.array([partyTotals,]*seats, ).T/divisors.astype(float)
    sortedPC = np.sort(np.ravel(pseudoCandidates))
    lastSeated = sortedPC[-seats]
    theSeated = np.where(pseudoCandidates >= lastSeated)
    partySeats = np.bincount(theSeated[0], minlength=len(partyTotals))
                                        # number of seats for each party
    inx = np.nonzero(partySeats)[0]     # only those with at least one seat
    seated = zip(inx, partyTotals[inx], divisors[partySeats[inx]-1])
                                        # parties with at least one seat,
                                        # number of votes that party got,
                                        # and divisor for last seated in
                                        # the party
    theNotSeated = np.where(pseudoCandidates < lastSeated)
    partyNotSeats = np.bincount(theNotSeated[0], minlength=len(partyTotals))
                                        # number of non-seats for each
                                        # party
    inx = np.nonzero(partyNotSeats)[0]
    notSeated = zip(inx, partyTotals[inx], divisors[partySeats[inx]])
```

USENIX Journal of Election Technology and Systems (JETS)
Volume 3, Number 1 • December 2014
www.usenix.org/jets/issues/0301

```
                                            # parties with at least one
                                            # unseated, number of votes
                                            # for the first non-seated
                                            # in the party
    if (lastSeated == sortedPC[-(seats+1)]):
        raise ValueError("Tied contest for the last seat!")
    else:
        return partySeats, seated, notSeated, lastSeated, pseudoCandidates

def uMax(win, lose):
    '''
    finds the upper bound u on the MICRO for the contest
    win and lose are lists of triples: [party, tally(party), divisor]
    the divisor for win is the largest divisor for any seat the party won
    the divisor for lose is the smallest divisor for any seat the party lost
    See Stark and Teague, 2014, equations 4 and 5.

    Input:
        win:  list of triples--party, tally(party), divisor
        lose: list of triples--party, tally(party), divisor

    Returns:
        maximum possible relative overstatement for any ballot
    '''
    u = 0.0
    for w in win:
        for ell in lose:
            if w[0] != ell[0]:
                u = max([u,
                  (float(ell[2]) + float(w[2]))/float(ell[2]*w[1] - w[2]*ell[1])])
    return u

def minSampleSize(ballots, u, gamma=0.95, alpha=0.1):
    '''
    find smallest sample size for risk-limit alpha, using cushion gamma \in (0,1)
    1/alpha = (gamma/(1-1/(ballots*u))+1-gamma)**n
    Input:
        ballots: number of ballots cast in the contest
        u:       upper bound on overstatement per ballot
        gamma:   hedge against finding a ballot that attains the upper bound.
                 Larger values give less protection
        alpha:   risk limit
    '''
    return math.ceil(math.log(1.0/alpha) /
                     math.log(gamma/(1.0-1.0/(ballots*u)) + 1.0 - gamma))

# final 2014 Danish EU Parliamentary election results from
#   http://www.dst.dk/valg/Valg1475795/valgopg/valgopgHL.htm
# there were two coalitions: (A,B,F) and (C,V)
# There were 13 seats to allocate.
```

```
#
#  Official results by party
#
A = 435245
B = 148949
C = 208262
F = 249305
I = 65480
N = 183724
O = 605889
V = 379840
Ballots = 2332217  # includes invalid and blank ballots
nSeats = 13  # seats to allocate
#
# allocate seats to coalitions
#
coalitionTotals = np.array([A+B+F, C+V, I, N, O])  # for coalitions
coalitionSeats, coalitionSeated, coalitionNotSeated, coalitionLastSeated,
                coalitionPCs
                = dHondt(coalitionTotals, nSeats, np.arange(1, nSeats+1))
print 'A+B+F, C+V, I, N, O:', coalitionSeats
#
# allocate seats within coalitions
#
nABFSeats = coalitionSeats[0]
nCVSeats = coalitionSeats[1]
ABFSeats, ABFSeated, ABFNotSeated, ABFLastSeated, ABFPCs
         = dHondt(np.array([A, B, F]), nABFSeats, np.arange(1, nABFSeats+1))
CVSeats, CVSeated, CVNotSeated, CVLastSeated, CVPCs
         = dHondt(np.array([C, V]), nCVSeats, np.arange(1, nCVSeats+1))
#
print 'A, B, F:', ABFSeats, ';  C, V:', CVSeats
#
ASeats = ABFSeats[0]
BSeats = ABFSeats[1]
CSeats = CVSeats[0]
FSeats = ABFSeats[2]
ISeats = coalitionSeats[2]
NSeats = coalitionSeats[3]
OSeats = coalitionSeats[4]
VSeats = CVSeats[1]
allSeats = [ASeats, BSeats, CSeats, FSeats, ISeats, NSeats, OSeats, VSeats]
print '---------------\nSeats to parties A, B, C, F, I, N, O, V: ', allSeats
print 'Seated coalitions, votes, divisor:', coalitionSeated
print 'Non-Seated coalitions, votes, divisor:', coalitionNotSeated
#
#  Set audit parameters
gamma = 0.95  # tuning constant in the Kaplan-Wald method
alpha = 0.001 # risk limit
#
```

```
u = uMax(coalitionSeated, coalitionNotSeated)
print Ballots*u
n = math.ceil(math.log(1.0/alpha) /
              math.log(gamma/(1.0-1.0/(Ballots*u)) + 1.0 - gamma))
print n
```