

**From:** section 47F  
**Sent:** Friday, 19 August 2022 5:08 PM  
**To:** section 47F  
**Cc:** section 47F  
**Subject:** RE: AEC party sampling methodology - summary of meeting of 21 July 2022 (LEX1984) [SEC=OFFICIAL]

Dear section 47F

Here are your options for question 1.

- A. Do not sample at all, check the list until you find 1500 eligible or until they can't make the required number.
- B. Sample from the list according to the previous advise documents and use the tool to control the two risks. A list of 4680 would as you point out require a sample size of 564 and max number of denials allowed to be 399, with 6% and 2% risks of false rejection and false acceptance.
- C. Take what the sampling theory call two phase sample. First sample from the larger list randomly and then sample from that sample. (In this case it is hard to see the point of this given the processes you apply. Unless you collect some auxiliary information from the first phase and then further (detailed) information in the second phase with a smaller sample. If you do this randomly you can control the risks but your (total) sample size still need to be in the vicinity of the one in B. (It all depends on what you can collect and use to optimise the second phase.)

Considering you have limited resources and no practises to use and apply auxiliary information in method C our advice is to use B. C makes sense only if you can acquire auxiliary information in the first phase that is useful to guide (decreasing) the sampling size in the second phase. My understanding is that such options are limited.

A practical view I would suggest to you is to take out a fee for parties who ask you to test a larger sample than practically affordable. As I said the White pages party can send you a copy of the White pages. With a very low proportion of eligible members you would have to investigate a very large sample to be reasonably sure they don't have at least 1500 members. Controlling the risks levels unaffordable to manage.

For the **second question** it is important to recognise that the probabilities / risks exist **before** you do something and the probabilistic interpretation are over several repeated events. It is also influenced on how you go about doing it in practise.

For A above the both risks are practically zero. At a (high) cost you will be able to find whether the list truly has >1500 or not.

For B it is <0.06 and <0.02 if you use the sample size above and mentioned in your observation (1) below. If you use another sample size the risks change all depending if the **true** value is above 1500 or not. A smaller sample size increases the risks.

For C it is basically still the same problem as B from a risk perspective should you not have any extra information collected before your second phase sample. Hypothetically, if your 1650 was a random first phase sample and you then sampled from this list you would still need to investigate 564 with 399 to achieve <.06 and <.02 risk rates.

As for your observations.

- 1. Is correct
- 2. If you have a list of 1650 (not case C above which is a population list of 4680) the table said sample size 60 and max denials 9. I checked the calculator and you are right 55 and 8 will also satisfy the constraints and it is better. These are anomalies appearing because the functions are discontinuous I think the table was



generated by a test algorithm rather than checking/optimising each case as a cheat sheet. You will have to ask [REDACTED] about it but it is not surprising and of minor practical importance.

3. It seems as you have explored the calculator. With N=4680 and n=43 you do get a about 29% and 29% for both cases. This assumes that the 1650 is a random sample from the 4680 and that this is done in advance as part of the testing procedures. Recall my comment above about risks for false conclusion. Another aspect with 29 and 29 is that the risk for **any** false conclusion increases and that is quite unorthodox.
4. Yes see comment in (3). It is the risk assessment you must consider. Lowering your appetite on one will increase the other and vice versa. I suppose you also have to have the same one for all applying parties.

All the best,

section 47F

---

**From:** section 47F  
**Sent:** Thursday, 18 August 2022 5:59 PM  
**To:** section 47F  
**Cc:** section 47F  
**Subject:** RE: AEC party sampling methodology - summary of meeting of 21 July 2022 (LEX1984) [SEC=OFFICIAL]

**CAUTION:** External email. Do not click links or open attachments unless you recognise the sender and know the content is safe.

---

Dear section 47F

### **Request for advice about ABS Testing Methodology**

Thank you for your previous advice in relation to the ABS testing methodology and testing of VoteFlux.Org | Upgrade Democracy! (the Party). The Electoral Commission has asked us to seek your advice on what options are available in the event the Commission wishes to conduct further testing in relation to the Party's list of 4680 members to establish whether the party has 1500 members.

Accordingly, we would be grateful for your advice on two matters:

1. **What different options are available to the AEC to test the list of the Party?**
2. **For each of those options what is the probability of falsely accepting a list? What is the probability of falsely rejecting a list?**

The Commission is meeting next Thursday 25 August 2022 to discuss this matter. However, if you need more time to provide advice the matter may be adjourned to a later date. We would be most grateful if you could respond by reply email at your convenience. We would also be happy to discuss with you further.

### **Observations of AEC Legal with respect to the questions posed**

In relation to the Commission's questions, we make the following observations, drawing on our understanding from previous discussions with you. In providing your response, you may wish to refer to these points and state whether you disagree with them.

Based on our previous discussion with you, it seems to us that the appropriate way to decide how to test the Party's list is to input values in to the ABS calculator in relation to list size and desired probabilities of false acceptance/rejection and then see the sample size needed to achieve those probabilities. The Commission would need to decide on a sample size and risk level pairing with which it was comfortable.

1. We have used the ABS calculator (attached), to calculate that sampling a list of **4680** members, where the desired probability of false rejection is **6%** and the desired probability of false acceptance



is 2% (the default parameters), would require the AEC to test an approximate sample size of 564 and a maximum number of denials allowed of 399.

2. We note that testing a list of 1650 members where the desired probability of false rejection is 6% and the desired probability of false acceptance is 2% requires an approximate sample size of 55 and a maximum number of denials of 8.
3. If the AEC were to try to use a similar sample size of around 50-60, then it seems to us that the risks of false acceptance and rejection would both increase. Using the calculator, we calculate that for testing a list of 4680 members, a sample size of 61 members with an approximate maximum number of denials of 43 can be used. This gives a desired probability of false rejection is 29% and the desired probability of false acceptance is 29%.
4. It would of course be possible to calculate different sample sizes in between these two ranges based on different probabilities of false acceptance and rejection.

section 47F

Legal Services Section | Legal & Procurement Branch  
Australian Electoral Commission

section 47F

---

From: section 47F

Sent: Friday, 29 July 2022 5:09 PM

To: section 47F

Cc: section 47F

Subject: RE: AEC party sampling methodology - summary of meeting of 21 July 2022 (LEX1984) [SEC=OFFICIAL]

Dear section 47F

I think you got it. I have added some extra clarification and some edits in red below. The purple I suggest you delete. (It is impossible to use statistical theory to say anything about the confidence of rejecting (or accepting) that there are  $\geq 1500$  among the 4680. You can be *pretty sure* that there are not 1500 among the 1650 but I cannot calculate how sure you can be and it is a bit beside the point. With a small sample size and poor quality list it's just not how you would set up and perform a statistical test.

(As a simple example, if the first 150 you checked were invalid you'd need to find one more among the remaining 1500 to definitely reject with 100 % certainty. At the same time if the first 1499 were members you still need one more of the remaining 151 to 100% accept. Through randomised sampling probability theory controls the risks and optimises (minimises) the required sample size for those controls. If there are systematic patterns in the lists and it is not randomised it is more or less impossible to be very accurate about the decision probabilities of a test.)

Hope this helps.

Best regards,

section 47F

---

From: section 47F

Sent: Friday, 29 July 2022 2:18 PM

To: section 47F

Cc: section 47F

Subject: AEC party sampling methodology - summary of meeting of 21 July 2022 (LEX1984) [SEC=OFFICIAL]

**CAUTION:** External email. Do not click links or open attachments unless you recognise the sender and know the content is safe.

---

Dear section 47F

Thank you for meeting with **section 47F** and I last week on Thursday 21 July. At that meeting we discussed the ABS methodology for sampling and testing membership of political parties used by the AEC. You kindly reviewed some questions we had and considered some issues raised by applications for review of the decision to deregister VoteFlux.Org | Upgrade Democracy! (the Party). The purpose of this email is to summarise the key points raised in that discussion. I would be grateful if you could confirm that I have set out your advice correctly. If not, I would be grateful if you could edit the response or reply clarify any points.

If possible, we would be grateful for your response by Monday 1 July 2022. We may provide this information to the Electoral Commission for the purposes of their review of the decision to deregister the Party.

### **1. What conclusions can be drawn in relation to the list of 4680 names in the Party's list, following the delegate's decision to test the top 1,650 names of that list (the sub-list)?**

With respect to the test conducted it is not the randomisation or not that is causing the false rejection (or false acceptance) rate to land outside the limits. It is the sample size that control the test conditions. If you had randomised you would have been able to say something about the whole list as the final sample (probabilistically) would have referred to the 4680 not just the 1650. Nevertheless you still would have needed a bigger sample size to get the desired risk rates.

You advised that as the Party's list was not randomised before the sub-list was made (as only the top 1650 names were selected) it is not possible to draw any meaningful statistical conclusions about the Party's **whole** list of 4680 from the results of testing the sub-list. **You can only say something about the selected 1650.** This is because, without randomisation there is *no chance* of the other records being selected. You explained this by the analogy of attempting to sample **from** a deck of cards for aces. Having failed to first shuffle the deck, **chunking the bottom half away, and then sample from what is left will not give you useful information about all original cards. Those thrown away were never in the running.**

**The test done indicates a list with low proportion of eligible members among the 1650.**

**DELETE** As would usually be the case, as the test of the sub-list failed, it is possible to conclude (to the level of confidence to which the test is set) that the sub-list of 1,650 did not include 1,500 members.

### **2. What are the alternative options for testing the larger list?**

You explained that a larger sample size would be required to test a large list to the same degree of confidence. Current sampling is calibrated so that the probability of accepting an invalid list is less than 2%, and the probability of falsely rejecting a valid list is less than 6%. You estimated that the sample size required to test the full list of 4680 names would be over 300. You also advised that excel spreadsheet calculator provided to the AEC could provide information about the sample size required to test a list of 4680 names to different levels of confidence.

We have now used the calculator (attached), to calculate that sampling a list of 4680 members, where the desired probability of false rejection is 6% and the desired probability of false acceptance is 2%, would require an approximate sample size of **564 and a maximum number of denials allowed of 399.**

### **3. Does filtering names affect the error rates?**

You explained that in your view, the arguments made in **S 47F** paper with respect to filtering names increasing the error rate are without foundation. You explained that, provided the filtering process is done in accordance with the ABS methodology, filtering names works in favour of parties by removing from a list members who would not have been capable of meeting the requirements.

**Correct this would improve the 'quality' of the list and decrease the occurrence of finding denials (non-members) in the list sample.**

#### 4. Your general comments on S 47F paper

You considered that it was not instructive to consider in depth S 47F hypothetical example. That example started from the premise that the party has more than 1,500 and sought to prove that a specific list could be rejected by the sampling methodology.

You agreed with the general proposition that if the correct sampling size was not adopted in relation to a larger list, the likelihood of false rejection increased. This is shown by the calculator.

We also discussed generally the rationale for requiring a smaller sampling size as a practical and fair method for testing party lists. We discussed the difficulties of testing a larger list. Since the requirement is minimum 1500 a party with a very large list that is 'low quality' in the sense that it contains a high percentage of non-members will require a very big sample size to control false rejection risks. This relates to the incentives of parties to keep good records of their members and provide the AEC with a high quality list. Providing large low quality lists should be discouraged.

section 47F

Legal Services Section | Legal & Procurement Branch  
Australian Electoral Commission

section 47F

#### DISCLAIMER:

If you have received this transmission in error please notify us immediately by return email and delete all copies. If this email or any attachments have been sent to you in error, that error does not constitute waiver of any confidentiality, privilege or copyright in respect of information in the email or attachments.

#### DISCLAIMER:

If you have received this transmission in error please notify us immediately by return email and delete all copies. If this email or any attachments have been sent to you in error, that error does not constitute waiver of any confidentiality, privilege or copyright in respect of information in the email or attachments.