

December 15, 2014

FROM: Chris Jerdonek, Vice President  
TO: Elections Commission;  
John Arntz, Director of Elections

SUBJECT: November 4, 2014 Election Observations (12 pages, updated)

This document contains personal observations and thoughts related to the San Francisco November 4, 2014 Consolidated General Election.

## 1. Summary

Overall I thought the election was conducted excellently. In particular, this election had a number of substantial improvements over the previous election in the areas of public transparency and auditing. The willingness and ability of Director Arntz and Department staff to take on and implement improvements like these in the span of only five months shows true organization and commitment.

In addition, on top of all of the usual election tasks, the Department had to overcome a number of significant obstacles, several of which were unanticipated. These include the lengthy 5-card ballot, thousands of vote-by-mail ballots getting damaged by the Post Office, the ballot printing error, and the brief server downtime on Election Night. Much credit should go to Director Arntz and his staff for gracefully handling these issues.

In response to various observations, this document also discusses a number of possibilities for improvement. Many of these are around the related areas of “open data,” transparency, and reporting, and could build on previous improvements.

## 2. Observation Activities

My in-person observation activities included the following:

- Attended experienced inspector training class at the San Francisco Public Library;
- Polling-place inspector on Election Day for Precinct 7845;
- Observed canvass at City Hall on Tuesday, November 11;
- Observed loading of VBM ballots at City Hall on Wednesday, November 12 for transport to the warehouse at Pier 48 B;
- Observed unloading of VBM ballots and canvass at the Pier 48 B warehouse on Wednesday, November 12;
- Observed the Alameda County 1% random selection of precincts on Thursday,

- November 13 in Oakland (for comparison purposes);
- Observed the 1% random selection of precincts at San Francisco City Hall on Monday, November 17;
- Observed the 1% manual tally at Pier 48 B on Wednesday, November 19.

### 3. Improvements over Previous Election

This election had a number of substantial improvements over the previous election held on June 3, 2014. This section lists some of these improvements that I observed.

#### 3.1 Public Outreach

- Public e-mail announcement list.** After the June election, the Department created an e-mail list that members of the public could subscribe to for public notices. This made it much easier for members of the public to learn about election-related activities.
- Public education via Twitter.** The Department did a great job of using photos on Twitter to educate the public about what goes on “behind the scenes.” Here is a link to the Department’s Twitter account: <https://twitter.com/SFElections>

#### 3.2 Election Auditing: Reporting

- Posting preliminary Statements of Vote in open and machine-readable format.** The Department posted the information in the preliminary Statements of Vote in an open and machine-readable format for the first time. Specifically, the format was tab-delimited (aka tab-separated value or TSV). The daily posting of these files began on Sunday, November 9. The TSV files included the same information that the PDF Preliminary Statements of Vote contained, including (1) the totals in each precinct for each contest, further broken into in-precinct and VBM subtotals; (2) undervote and overvote totals for each contest; and (3) district and neighborhood totals for each contest. Machine-readable files like these make it much easier for members of the public to analyze the data, thus increasing election transparency.
- Posting raw WinEDS results files.** In addition to the TSV files, the Department also posted the raw results files exported from the Department’s WinEDS election software. These are text files like the TSV files, though they are more difficult to work with. The posting of these files began on Thursday, November 6. These files were especially useful in the three days prior to releasing the TSV files because these were the only files available that contained precinct-level data in a machine-readable format.

#### 3.3 Election Auditing: 1% random selection and manual tally

- Including all ballots in the 1% manual tally.** In this election, all ballots were included in the 1% manual tally, allowing for a true audit. In prior elections, ballots tabulated

- after Election Day were not necessarily subject to the 1% manual tally.
2. **Counting all ballots before the 1% random selection.** In addition to including all ballots in the 1% manual tally, the Department also waited until all ballots were counted before randomly selecting the precincts for the 1% manual tally. This decision required a significant amount of extra work on the part of the Department, because the Department needed to sort all VBM ballots counted after Election Day by precinct, and then transport them to the warehouse at Pier 48 B, all while maintaining a proper chain of custody. It needed to do this to allow the retrieval of ballots for the precincts selected by the 1% random selection. The Department should be commended on how it planned for and carried out this sorting and transportation process. This was a crucial step towards the 1% manual tally being a true public audit because it allowed the precinct-level totals that would be manually checked to be posted publicly before randomly selecting the precincts for checking. In particular, the vote totals were not still being added to or altered between the random selection and manual tally (though I did observe rare exceptions to this related to RCV write-ins).
  3. **Using 10-sided dice for the 1% random selection.** For the first time, the Department rolled dice for the 1% random selection of precincts. Specifically, the Department rolled three translucent 10-sided dice colored red, white (“clear”), and blue to randomly choose numbers in the range 000, 001, 002, and so on up to 999. The 594 precincts corresponded to the numbers in the range 001 and 594. The dice rolling was done publicly and video-streamed live to YouTube. The procedures were explained to the public clearly beforehand (both in writing and orally). The selected precincts were announced afterwards via a press release that was posted on the Department website and e-mailed to the public e-mail list. The process of rolling dice was much faster and more clearly random than the previous procedure of choosing cards from a box. Rolling dice is considered a best practice by election security experts. Here is a short 40-second video clip of this election's random selection that I recorded:
    - <https://www.youtube.com/watch?v=URzMr1tfQtE>

### 3.4. Elections Commission

- **Forwarding security plan to Board of Supervisors.** The Elections Commission forwarded the Sheriff's transportation and security plan to the Board of Supervisors (as required by Sec. 13.104.5 of the San Francisco Charter).

## 4. Possibilities for Future Improvement

This section discusses possible areas of improvement for future elections inspired by specific observations. The size of this section should not be taken as a negative indication. Rather, even the best processes have opportunities for improvement.

## 4.1. Election Commission

- **Employee Waiver (repeat from June 3, 2014).** The employee waiver that Director Arntz requested and that the Commission approved at its August 20, 2014 meeting does not appear to have been approved by the Board of Supervisors, as required by Sec. 13.104.5 of the San Francisco Charter. Such a motion is required for non-Department City employees to provide election-related services to the Department. The Board agendas between August 20 and November 18 do not seem to contain such an agenda item. For the June 3, 2014 election, the Elections Commission did get such a motion passed by the Board, though not until one week after Election Day on June 10, 2014.

## 4.2. Election Data and Results Reporting

This section relates to election data and results reporting and is divided into two subsections. The first section lists some things I observed that present opportunities for improvement. The second section describes some ideas to address the observations in the first section.

### 4.2.1. Election Data: Areas for Improvement

- **Website stability.** The Department website seemed to have some stability issues on election night that made it harder for members of the public to access election results. For example, SFGate published an article about this on November 5.
- **Easy access to contest information.** One member of the public explained to me that it was not easy for him to get official information about what is on the ballot aside from downloading and reading through the 200+ page Voter Information Pamphlet (VIP) PDF posted on the Department website. This makes it harder to find and access information about specific contests. This also makes it harder to access contest information in different formats and in more digestible forms (e.g. short summaries). Specific contest information that would be useful to make more accessible includes things like—
  - Candidate names and statements,
  - Proposition digests prepared by the Ballot Simplification Committee, and
  - Proposition arguments for and against.
- **Publicizing in advance the availability of machine-readable results.** The availability of machine-readable election results was not publicized in advance. Releasing election results in a machine-readable format is useful because it lets members of the public use computer programs to process the data automatically, which in turn can benefit the broader public in various ways. For example, a media organization could automatically generate visualizations of the election results for the public whenever the Department posts new results. However, such computer programs can be sophisticated and need to be prepared and written in advance. This can take individuals or organizations days or weeks. If the availability of the data is not

- known advance, then members of the public cannot plan for this.
- **Documenting the format of machine-readable results in advance.** Related to the above, if the format of the machine-readable data is not described and documented in advance, then individuals cannot finish writing computer programs in advance – even if they know machine-readable data will eventually be released. They would need to wait until after Election Day to discover the format of the data, and only then can the computer programs to download and process that data be written. Thus, it would be helpful for the public to know in advance both what information will be released as well as the specific format of that data (for example including sample data files they can prepare and test against).
  - **Availability of machine-readable data on election night.** While it was very positive that the Department released election results in an open and machine-readable format, this did not happen until much later. The harder-to-use WinEDS “raw” files were released starting two days after the election on Thursday, November 6. The nicer TSV files were released starting five days after the election on Sunday, November 9. The highest level of interest is usually on election night and the day after, so it would be useful if the machine-readable data were also available then (or more generally, made available at the same time more human-readable data is made available).
  - **Reporting ballots remaining.** At least one member of the public told me that it was difficult or not possible to ascertain details about how many ballots were left to be counted from day to day, including what type (e.g. VBM or provisional) and from what precincts. This type of information can be useful in determining the likelihood of any outcomes changing as more ballots are counted (since different precincts tend to favor different candidates at different rates). This is especially useful in closer races, like the District 17 State Assembly race in this election.
  - **Understanding changes between preliminary reports.** Similar to the previous point, especially in the first few days of counting, it is not easy for members of the public to know which precincts contributed to changes in election totals from day to day and by how much (e.g. how many ballots were added to each precinct and what type of ballot, etc). This information can also be useful in determining the trajectory of later counts, and specifically whether any outcomes are likely to change. Once the TSV files started being released, this type of information could at least be calculated, for example by using a computer program to subtract from one day the TSV totals of the previous day. One caveat is that this method requires access to the TSV files from previous days. Since the Department website only displayed the most current TSV file, members of the public would have needed to manually save the TSV file each day, and record the date that file was generated.

#### 4.2.2. Election Data: Ideas for future

Many of the observations listed above can be addressed by embracing and implementing a policy of open data. “Open data” is a philosophy of providing data to the public not just in human-readable forms, but also in ways that are optimized for computer and software use. For example, this can include exposing the data in machine-readable, open formats (like JSON, XML, CSV, TSV, etc) and making that data available in ways that are easy to discover

and download automatically using computer programs (e.g. using web APIs, also known as web services). Informally, these can also be described as “developer-friendly” ways.

Developer-friendly policies benefit not just software developers but also the general public. The reason is that information and software products created using open data are often intended for the general public (e.g. news analyses and consumer “apps”). To provide a couple examples, if detailed contest information were released in machine-readable forms, then—

1. Media organizations could create different “views” of the contest information on their own sites using automated tools, which would provide more opportunities for educating the public about the election.
2. App developers could create digital ballot worksheets like the paper version that appears in the last couple pages of the paper VIP (and pages 20 and 21 in the electronic PDF version). People use such worksheets to help them remember how to vote when entering a polling place.

In the absence of an API, information like the above can be onerous and time-intensive to collect. For example, it could require manual data entry; manually cutting, pasting, and reformatting contest information from the Voter Information Pamphlet (VIP); or writing custom software to “scrape” the Department website.

To give an example, Project Vote Smart is a national database of contest information that exposes this information through an API called the Vote Smart API ( <http://votesmart.org/share/api> ). However, to build this API each election, Project Vote Smart needs to employ the efforts of thousands of student interns and volunteers to do repetitive data entry from various election sources. The League of Women Voters “Smart Voter” Project ( <http://classic.smartvoter.org/> ) needs to conduct similar measures. If Departments like San Francisco exposed their contest information in an open machine-readable form, then the process of collecting election information could be automated and hence be done more cheaply, accurately, and efficiently.

In addition to organizations like Project Vote Smart and the League of Women Voters, other individuals and organizations likely to take advantage of open data include—

- Media organizations;
- Academic and nonprofit researchers; and
- Software companies and app developers, including civic-tech groups like Code for America (which also has a San Francisco branch).

Here are some more specific suggestions to address the needs listed in the “Election Data: Areas of Improvement” section above:

1. **Leverage a CDN.** A content delivery network (CDN) is a generic term for a third-party service specialized in serving static web sites. They are designed to be affordable and to handle spikes in website usage without system downtime. They can handle spikes

in usage because CDN's serve data from multiple redundant servers. They are low-cost because they are limited to static content (i.e. content that is the same for different users). This makes it ideal for serving static HTML pages like human-readable election results, while simultaneously avoiding server downtime during periods of high usage like election night.

2. **Leverage DataSF.** DataSF is a citywide open data portal launched by the City and County of San Francisco in 2010. As a citywide portal, it focuses on making data available across San Francisco departments in a common citywide format. By leveraging DataSF, the Department of Elections could potentially free up staff time and avoid having to create and maintain their own data API's. Also, my understanding is that the San Francisco data portal team can assist departments in leveraging the data portal. Specifically, they can write automated scripts to transfer data sets from Department servers to the data portal servers on a regular schedule. Lastly, because DataSF manages their own IT infrastructure and can benefit from economies of scale, the Department of Elections could rely on DataSF for avoiding API downtime.
3. **Expose basic election data sets through an API.** Here are some examples of data sets that the Department could make available through a data portal like DataSF:
  - Contest information like—
    - Candidate names, statements, and affiliations (including the Chinese spellings of names),
    - Proposition digests prepared by the Ballot Simplification Committee,
    - Proposition arguments for and against,
    - Information about contests in multiple languages,
    - The current, up-to-date listing of qualified write-in candidates,
  - Precinct information like—
    - What neighborhood and district each precinct belongs to,
    - The address of the polling place in the precinct,
    - Whether the precinct is all-mail,
    - The precinct name and ID,
    - The geographic data of each precinct (census blocks, etc),
  - Election results like—
    - Time-stamped preliminary Statement of Vote data,
    - Ballots remaining to be counted by precinct and type,
    - RCV ballot image data.
4. **Maintain a “developer” page on the website.** A developer page could be added to the Department website – serving as one-stop shop documenting what election-related information is available in developer-friendly formats. The page could document what APIs are available and where they are. The page could also document any data file formats for machine-readable data released by the Department (e.g. RCV ballot image file formats, WinEDS raw file formats, TSV file formats, etc), especially if such data is not part of an API. The page could also link to any open source software used or recommended by the Department.

### 4.3. Ranked Choice Voting (RCV)

This section lists some areas of improvement related to RCV.

- **Election night reporting.** Members of the public stated that in 2012, the Department released RCV round-by-round reports and ballot image files with both the first and last reported election totals on election night. This election, from my observations, the RCV report was generated only with the last report (sometime after 1AM, perhaps as late as 2AM), and the ballot image reports sometime the following morning. It would be good if the Department release the round-by-round and ballot image reports both with the first and last election night reports.
- **Final-round reporting.** The RCV Reports on the Department website do not always continue to two candidates. This prevents voters from being able to know the “mandate” of the winner. For example, in the D10 Board of Supervisors race, the final round of the report has three candidates and Malia Cohen winning with 51.7%. If the report carried the elimination to two candidates, then voters would have been able to see that Malia Cohen won 64.1% against Tony Kelly in the final round. A letter to the editor was also published about this issue in the San Jose Mercury News on November 12, but about the Oakland Mayor's race (titled “Schaaf's 'mandate' claims exaggerated”). The letter writer confused the percentage in the final 3-candidate round with the mandate that results from continuing to two candidates. It would be worth considering always reducing to two to provide more information to voters.
- **Ballot images files for all RCV races.** On most days after the election, the ballot image file contained only the ballots for the D10 Supervisor race, even though there were other RCV races on the ballot. However, the weekend of November 8, the ballot image files contained the ballots for the other RCV races as well. For greater transparency, the Department could release the ballot image data for all RCV contests, whether or not rounds of elimination were necessary.
- **More developer-friendly ballot image format.** The current ballot image file format is not as easy to work with as it could be. Currently, the data format is native to the proprietary WinEDS system. It would be nice if the data format were more standardized and easier to parse. This could be solved if the data were released through a data portal (as described above) rather than simply making the raw files available.
- **More substantive RCV education.** Currently, much of the Department's RCV education focuses on the “how” of voting rather than the “why.” For example, voters are told not to rank the same candidate three times, but they are not told the reasoning behind that guidance. In particular, they are not told why it helps to rank a different candidate second. If voters were given more substantive education, perhaps we would see more voters using their rankings for the fullest benefit.

### 4.4. 1% Random Selection

The processes around the 1% random selection of precincts are already very good in my

opinion, but here are some ideas for improving the process further.

- **Roll the dice in a shallow box or tray.** Currently, the dice are rolled on the surface of the front desk, which is somewhat narrow. Consequently, the dice must be rolled gently to prevent them from bouncing onto the floor. If the dice were rolled in a shallow box or tray (i.e. something with at least three “sides”), then the dice could be rolled more forcefully. This would allow the dice to bounce around more before landing on three numbers.
- **Note the contests for the “extra” precincts.** For the precincts beyond the initial one percent, it would help to note on the large paper next to the precinct name the contests that each additional precinct is being selected for (e.g. “Supervisor D6,” etc).
- **Let members of the public roll.** For increased transparency and public participation, it might be good to let a variety of people roll the dice (e.g. Department staff and members of the public).

#### 4.5. Website Navigability

The Department website feels harder to navigate than it should. For example, the home page feels cluttered with many links, and it is not always clear where certain information is located. In some cases, more than one page has the same or similar information, so it is not clear if you are on the “right” one. To provide one example, during the election there were at least two pages about voting: a permanent (or “evergreen”) page about voting, and a page about voting in the current election.

Some other observations:

- A friend of mine thought “Candidates and Campaigns” was the link to information about what contests were on the ballot. Instead, it is the link for campaign participants.
- The “Election Results and Archives” hierarchy seems to have more levels of nesting than it needs to. It would be more convenient if “Election Archives by Year” linked directly to all of the elections instead of listing all of the elections but only providing links to individual years. The separate “year” pages seem unnecessary.
- Related to this, there are cases where information about a given past election are in completely different sections of the website. In particular, the “Candidates & Campaigns Archive” (<http://sfgov2.org/index.aspx?page=2207>) is a list of past elections completely separate from the main “Election Archives by Year” (<http://sfgov2.org/index.aspx?page=1671>). For the purposes of researching past elections, it would be good if each past election had a single entry point. This entry point could link to all the information about the election (e.g. the candidate listings, VIP, as well as results) instead of spreading this information across separate sections of the site.
- Currently, navigating to the home page from anywhere on the website is not as intuitive as it could be because you have to click on the word “Welcome,” which on many websites suggests instructions for newcomers to a website. More common home-page link text is “Home.”

Here are some further ideas to consider:

- Avoid having multiple links on the home page that point to the same information or page. This will reduce the number of links and may simplify the home page.
- During elections, feature prominent links on the “evergreen” pages that direct the website visitor to the corresponding page for the current election. This way visitors will be directed to the right page if they accidentally go astray.
- Consider presenting website visitors with fewer choices rather than more. For example, website visitors can be presented with more general options like (1) “For Voters” and (2) “For Campaigners” to be directed to main web pages specific to those audiences.

#### 4.6. Election Materials, Processes & Transparency

- **Election Materials.** Currently, some of the main education pieces like the postcard sent to all households in San Francisco have three languages intermixed throughout the piece. This makes the materials a bit harder to read for everyone. Since voters already have to sift through so much information, it would help if the materials can be simpler. Moreover, with the addition of a fourth language, education pieces with four languages would get even more complicated. Thus, it would be worth considering creating education pieces with only one or two languages at a time. This would allow for a cleaner, simpler, and more understandable design. It would also allow more information to be added to such piece while reducing overall complexity.
- **Election Plan and Processes.** This idea relates to two things: simplifying the Election Plan and increasing visibility and transparency around election processes. The latest Election Plan was 36 pages long with a 9-page appendix called, “Procedures for Processing Write-in Votes.” The Department could consider creating a “Rules and Procedures” page on the Department website (perhaps in a section related to observing) that contains documents spelling out various procedures. The page could start out with documents that already exist. In addition, sections of the Election Plan spelling out detailed procedures that remain the same from election to election could be moved from the Election Plan to this new web page. For example, the 9-page appendix on “Procedures for Processing Write-in Votes” could be moved to this page. Another possibility are the four or five pages spelling out in detail the processes governing remakes. This would simplify the Election Plan and change the focus in the Plan to things that change from election to election. Other possibilities for inclusion on a “Procedures” web page (not from the Election Plan) include—
  - RCV batch elimination rules
  - RCV tie-break rules
  - 1% Random Selection Procedures
- **Election Plan Glossary.** Adding a glossary is another possibility for improving the organization and readability of the Election Plan. When I first read the Plan, there were several terms and acronyms I didn't know the meaning of, and it wasn't easy to find out where in the Plan they were defined.
- **Metrics over time.** There are many valuable metrics that can be used to evaluate

different aspects of an election. Voter turnout is the most common such metric. Voter registration is another. Other metrics are the percentage of challenged VBM ballots and the percentage of Edge votes cast, both of which the Elections Commission already looks at. It would be useful to see how numbers like these change over time. Examples of quantities that would be useful to graph election over election include—

- Voter turnout
- Percent VBM voters
- Percent provisional voters
- Percent challenged VBM
- Percent language preference
- Voter registration

#### 4.7. Poll-working

- **Checkboxes in security procedures.** In the Inspector workbook, it seems like lots of pages in the opening and closing procedures didn't have checkboxes next to each step where they possibly did before in previous elections. Having checkboxes next to each step makes it easier to keep track of what you have done so far. In particular, checkboxes can help prevent steps from being skipped.
- **Poll-worker contact info.** My precinct staffing sheet did not list a phone number for one of my polling-place clerks, so I was unable to call them. Moreover, when I called the Department to ask how to proceed, the Department did not have a phone number on record for the clerk either. The Department only had an e-mail address. However, when I tried e-mailing the person, I did not receive a reply. It would be good to ensure that poll workers have a phone number when they are signed up. Otherwise, inspectors cannot call their clerks before Election Day as instructed.
- **Poll-worker training around RCV.** The trainer did a good job of explaining in detail how to vote in an RCV contest. However, the trainer went on to say that RCV is “highly contentious” and then proceeded to point out someone in the classroom that said they didn't like ranked-choice voting. This communicated what I felt was an unnecessarily negative tone. Alternatively, trainers could take a more positive tone by building confidence in current procedures (e.g. by discussing the benefits of current procedures). For example, in the case of RCV, one could explain why we use it and state some of its benefits.

#### 4.8. New Kiosks

In my polling place, the new kiosk created some issues throughout the day. First, my polling place clerk had difficulty making the kiosk stand up. When I investigated, I discovered that the rubber bands inside one of the three vertical chambers were not attached. I had to reach my hand through the seams of the chamber to reattach several of the rubber bands. One of the rubber bands could not reattach because the cardboard inside was bent. I don't know when they first became unattached.

Also, during the course of the day, the wind blew the kiosk over five or six times. One time, the kiosk fell or almost fell onto someone. This was mainly because of the location of the kiosk. Because our polling place was in a small garage, the kiosk had to be placed outside near the entrance to the driveway (as in previous elections). Because the kiosk was tall, it easily caught the wind. There was no obvious solution. The kiosk needed to be stabilized more at the top rather than anchored at the base.

#### 4.9. Miscellaneous

- **Proposition L wording.** A couple people commented to me that the ballot question for Proposition L was vague and potentially misleading. It read, “Shall it be City policy to change parking and transportation priorities?” In particular, the question said only that it should be changed and didn't say how or in what way.
- **Holiday phone message.** A friend of mine relayed to me that their friend tried calling the Department on Monday, October 13 to inquire about their VBM ballot not arriving. This was Columbus Day and the same weekend that some VBM envelopes were reported damaged. The person said they waited on hold for over an hour before giving up. They later realized that the Department was closed for the holiday. However, the Department phone message didn't note the holiday closure. The person was able to reach the Department the following day. It would be good to ensure that phone messages are updated during Department closures.
- **Loading ballots at City Hall.** I observed the loading of counted VBM and provisional ballots at City Hall for transport to the warehouse at Pier 48 B on Wednesday, November 12. The ballots were loaded into a truck. The ballots were transported in rolling carts that each contained up to 48 boxes of ballots, with each box containing up to around 500 ballots. This works out to 24,000 ballot cards per cart. When I observed, the truck at City Hall was loaded with about twelve carts of ballots. Unlike at the warehouse, no forklift was available at City Hall (due to the City Hall forklift being broken, apparently). The lift attached to the truck was also not aligned with the ground. The loading process appeared difficult without a forklift, and I was told that one of the carts tipped over during the loading process on an earlier occasion.