Terry W. Reilly

TO: Elections Commission

Date: 4/21/10

From: Terry Reilly

Number of pages including cover: 5

Regarding: MIT Analysis of Ballot Marking errors in SF's 2004 RCV election and Neely & Cook's "Whose Votes Count? Undervotes, Overvotes and Ranking in San Francisco's Instant Runoff Elections"

Dear Election Commission,

Data from report "Ballot Marking Errors in the first San Francisco Instant Runoff Election" shows differences of percentage of incorrect ballots in different Districts.

A district with per capita Income of >$75,000, 4% hispanics and 13% Asian population shows an incorrect ballot rate of 1.35%.

A district with per capita Income of ~$19,000, 26% hispanics and 46% Asian population shows an incorrect ballot rate of 2.69%.

The ballot marking error legend shows V for a valid vote, O for two more more votes in a column, and U for a skipped column. The columns are listed as 1st, 2nd, 3rd votes.

Neely and Cook confirm in "Whose Votes Count? Undervotes, Overvotes and Ranking in San Francisco's Instant Runoff Elections" for America Politics Research that voters had trouble navigating the IRV ballot and this confusion does not subside with more exposure to the IRV system. They note their research "implies voters might continue to overvote at similarly high rates at future IRV elections." Additionally, they point out the strong correlation of overvoting in minority and immigrant residents.

Sincerely yours,

Terry Reilly

If you make more money and are white, you are 50% less likely to make a mistake on your IRV ballot. Average Incorrect Ballot for IRV in SF 1.88%. Typical incorrect ballot for conventional elections is <0.1%. Data From "Ballot-Marking Errors in the first San Francisco Instant Runoff Election authored by Greg Dennis, MIT. Data from SF Dept. of Elections.

Table D2: Supervisor District 2 IRV Ballot Statistics

Legend		# Ballots	% Ballots
	Total	37796	100.00%
A single vote for in all three	**VVV**	18568	49.13%
2 or more votes in 3rd	**VVO**	13	0.03%
1 vote in 1st and 1 vote in	**VVU**	3600	9.52%
2 or more votes in 2nd	**VOU**	72	0.19%
1 vote in 1st & 3rd	**VUV**	209	0.55%
1 vote in 1st & 2 votes in 3rd	**VUO**	6	0.02%
1 vote in 1st	**VUU**	10648	28.17%
2 or more votes in 1st	**OUU**	73	0.19%
1 vote in 2nd & 1 vote in 3rd	**UVV**	17	0.04%
1 vote in 2nd & 2 votes in 3rd	**UVO**	1	0.00%
1 vote in 2nd	**UVU**	79	0.21%
2 or more votes in 2nd	**UOU**	3	0.01%
1 vote in 3rd	**UUV**	36	0.10%
2 or more votes in 3rd	**UUO**	2	0.01%
1 vote in 1st & 2 votes in 3rd	**UUU**	4469	11.82%

Number of Candidates:	5
Winner's Percentage of First Choices:	61.25%
Average Valid Votes Per Ballot:	1.96
Percentage Correct Ballots:	98.65%
Percentage Incorrect Ballots:	1.35%
Percentage Ballots with Overvotes:	0.45%
Percentage Ballots with Undervotes:	50.84%
Percentage Ballots with Skips:	0.93%
Per Capita Income:	$75,877.00
Percentage Hispanic:	4%
Percentage Asian:	13%

Table D11: Supervisor District 11 IRV Ballot Statistics

	# Ballots	% Ballots
Total	24375	100.00%
VVV	16312	66.92%
VVO	35	0.14%
VVU	1960	8.04%
VOU	61	0.25%
VUV	244	1.00%
VUO	13	0.05%
VUU	4039	16.57%
OUU	171	0.70%
UVV	18	0.07%
UVO	1	0.00%
UVU	66	0.27%
UOU	5	0.02%
UUV	40	0.16%
UUO	2	0.01%
UUU	1408	5.78%

Number of Candidates:	8
Winner's Percentage of First-Choices:	32.24%
Average Valid Votes Per Ballot:	2.37
Percentage Correct Ballots:	97.31%
Percentage Incorrect Ballots:	2.69%
Percentage Ballots with Overvotes:	1.18%
Percentage Ballots with Undervotes:	32.94%
Percentage Ballots with Skips:	1.60%
Per Capita Income:	$19,176.00
Percentage Hispanic:	26%
Percentage Asian:	46%

Again, the contextual factors operated as expected. Voters were more likely to use the ballot to its fullest in contests where more candidates ran, and that influence was relatively strong. They were also more likely to rank three candidates in precincts where more campaign money was spent. Last, the evidence here shows that familiarity with the IRV ballot does matter in regard to ranking candidates. The number of ballots on which three candidates were ranked was 7% to 9% higher in precincts where voters used it for the second time compared to those using it for the first time.

Discussion

We posed the question, "Whose votes count?" out of concern that the IRV reform might advantage some voters over others. In regard to overvotes, that concern appears valid. But the overall findings are mixed, with tendencies to undervote and rank candidates carrying more positive implications. To be clear, the data allow more rigorous tests of voting behavior in IRV elections than a comparison between IRV and plurality or TRS systems. Because those comparisons rely on precious few observations, the inferences they yield must be considered suggestive. When coupled with the more robust findings from the regression estimates and the natural experiment, however, we obtain some instructive insights in this first close analysis of IRV in the United States.

First, consider an important type of uncounted vote—overvotes. These are ballots that voters, in all likelihood, have marked in good faith and expect to count toward their desired outcome. Although they make up a fraction of all uncounted votes, they represent a significant failure of the voting system, the voter, or both. Although San Francisco's precinct-level optical scanners with error-correction indicators should produce relatively low rates of overvoting (Kimball & Kropf, 2005), the rates under IRV were as high as those in punch-card elections for similar offices (Sinclair & Alvarez, 2004). Note that the number of overvotes would be even higher if San Francisco's definition of spoiled ballots were less forgiving (see Note 5). It appears that voters had trouble navigating the more complex IRV ballot. This is supported by the fact that the number of candidates listed on the ballot was the strongest determinant of overvotes.

In addition, the null result from the natural experiment suggests that confusion with the IRV ballot might persist over time. We expected that practice with IRV would lead to fewer errors, but it did not. Although one might suspect that our measure of exposure was inadequate—that is, that one prior use

Downloaded from http://apr.sagepub.com by Terry Reilly on October 10, 2009

of IRV might be insufficient for learning to occur—voters' ranking behavior *did* change significantly after just one exposure. This implies that voters might continue to overvote at similarly high rates in future IRV elections. If additional research were to confirm this, it would identify an important shortcoming of the IRV reform, important because the proportion of African American and foreign-born residents in precincts was consistently and strongly related to the count of overvotes. And although those differences might not be a unique function of IRV, if IRV increases overvoting, then the reform serves to exacerbate those discrepancies. Given the import in equality of voice among voters, these findings deserve more study.

Although overvotes were more common, it appears that the IRV system produced relatively fewer undervotes when compared to previous San Francisco TRS elections. Such reductions are desirable and serve to equalize the expression of preferences across types of voters. We hypothesized that more, not fewer, undervotes would result from the added informational demands of IRV. This surprising overall result suggests that information costs did not present a significant barrier to voting under IRV. The more specific tests within the IRV elections also show that concerns about voters' language and verbal skills or the novelty of IRV working as a barrier to the vote were unwarranted. However, because the two contextual factors—campaign spending and the number of candidates running—were relatively strong predictors of undervotes, the availability of information about candidates *does* appear to affect behavior. We can only speculate at this juncture that other unspecified aspects of IRV counterbalance the additional information costs and see this as another topic worthy of further research.

We asked a second question, "Whose votes count more?" to address the possibility that some voters more than others would exercise their option to rank candidates. Because voters submitting fully ranked ballots provide more input than others, it matters whether or not those tendencies are distributed evenly across types of voters, especially those considered more at risk (e.g., the less educated and poorer). It is good news for democrats that the demographic influences on ranking were mostly absent and relatively small. The theory of information costs best explains tendencies to fully rank the IRV ballot, with more campaign spending and longer lists of candidates among the leading influences. As mentioned above, prior exposure to IRV also increased the likelihood of ranking three candidates, suggesting that experience and familiarity matters.

In sum, our findings suggest that voters adapted fairly well to IRV and that various types of voters experienced it in largely similar ways. The differences we observed in the types of voters who undervote match what occurs

Downloaded from http://apr.sagepub.com by Terry Reilly on October 10, 2009

under plurality and TRS systems, and ranking behavior was unrelated to most demographic measures. To be sure, the tendency for some voters to overvote more than others presents a problem. Further outreach to those communities could reduce these discrepancies in future elections, as could more rigorous training of precinct workers to help voters better understand how to correct ballot errors. We also see value in San Francisco's policy that forgives certain types of mistakes when marking the ballots. Clearly, this led to fewer spoiled ballots.

Local electoral reforms like San Francisco's move to IRV provide a useful lens through which to examine voting behavior. Although often perceived as rare, such experimentation is "alive and well at the sub-national level" in the United States (Bowler & Donovan, 2006, p. 2). This case helps us understand how voters adapt to more complex systems and sheds light on the influences in local elections, an arguably understudied venue. Although we have focused on voters' behavior as they marked their ballots, we note that a comprehensive assessment of IRV will balance our findings with other aspects of those elections, including turnout, representation, the nature of campaigns, the cost of election administration, and the perceived legitimacy of the system.

Notes

1. It is better known in comparative studies of election systems as the alternative vote.

2. The charter of the City and County of San Francisco (2008; which are consolidated) was amended to adopt IRV with Proposition A in March 2002. Originally called instant-runoff voting (IRV) and later dubbed ranked-choice voting, it is used to elect local officials, all of whom were previously elected in a two-round system.

3. Currently, five cities in the United States use IRV for local offices (Burlington, VT; Cary, NC; Hendersonville, NC; Tacoma Park, MD; and San Francisco), and three states use it in their overseas absentee balloting (Arkansas, Louisiana, and South Carolina). Other jurisdictions that have enacted IRV provisions but have yet to conduct IRV elections include Aspen, CO; Minneapolis, MN; Oakland, CA; Pierce County, WA; Berkeley, CA; and Ferndale, MI.

4. The city may limit the number of candidates voters rank based on the capacities of voting machinery, providing at least three can be ranked (City/County Charter Section 13.102[b]).

5. Two types of errors are allowed: If Column 1 contains no mark and Column 2 contains a valid mark, then the ballot is "advanced" and the mark in Column 2 is considered the voter's first preference. Also, if a voter chooses the same candidate in both Columns 1 and 2, then the mark in Column 1 is counted and Column 2 is ignored.

6. Rolloff is also widely studied. In theory, it is an undervote occurring when an item on the ballot is left blank. In practice, the measure usually includes both undervotes and overvotes.

7. In this system, voters gain no tactical advantage by ranking fewer than three candidates. Multiple votes for a single candidate are ignored (see Note 5). However, we note that some

Downloaded from http://apr.sagepub.com by Terry Reilly on October 10, 2009