A Measure to Detect Between-Candidate Relative Inequality Generated by Optical Scan Counters: An Analysis of the 2012 Presidential Election Data in South Korea

HeeKyoung Chun, Department of Epidemiology, Georgia Southern University, PO Box 8015, Statesboro, USA.1-513-410-1608 hchun@georgiasouthern.edu

Pierre-Jerome Bergeron, Department of Mathematics and Statistics, University of Ottawa, Canada. pierrejerome@gmail.com

HyunSeung Kim, Project BOO Inc. 268 Chungjeongro 3rd St., Seodaemungu, Seoul, South Korea.82-10-3096-3628 n2mart@gmail.com

OuJoon Kim, Project BOO Inc. 268 Chungjeongro 3rd St., Seodaemungu, Seoul, South Korea. 82-10-5218-3395 oujoon.k@gmail.com

Hwashin Hyun Shin*, Department of Mathematics and Statistics, Queen's University, 48 University Ave. Kingston, ON. Canada, K7L 3N6. 1-613-741-0117 hhshin@mast.queensu.ca

*Corresponding author

Department of Mathematics and Statistics, Queen's University,

48 University Ave. Kingston, ON.

Canada, K7L 3N6

613-741-0117

Email: hhshin@mast.queensu.ca

1. Introduction and Motivation

Many countries use electronic voting systems and such systems have shown result-changing errors through problems with software, hardware and procedures [1,14]. Appel, the Princeton group and others in cybersecurity and statistics have insisted that scanners and tally software are programmable and thus can be hacked [1,14]. Researchers discovered that the scanners had been misprogrammed or miscalibrated in some places [1,8,9,14]. Post-election vote-tabulation audits raised statistical issues [1,8,9]. The Monte Carlo simulation of Kobak et al. [9] confirmed high statistical significance of the observed phenomenon and its human-made nature.

For fast vote counting in elections optical scan (op-scan) counters in particular have been used to read, sort and count marked paper ballots. There are still manipulation issues that prevent op-scan counters from being an effective voting system: op-scan counters are known to be vulnerable to internal or external manipulations [1,15]. Limited studies have reported how electronic vote counting system can corrupt voting results.

According to the National Election Commission (NEC) of South Korea, presidential election in South Korea has used the op-scan counters for marked paper ballots since 2002 [5,10]. The op-scan counters first sort out the paper ballots into two groups: **classified** (sorted to each candidate by the op-scanners) versus **unclassified** (unsorted first by the op-scanners but sorted later manually by counting officials which include of government officials, teachers, commissioners, etc. [11]). Since the op-scan counters are the main tools for vote counting, it is necessary to examine the unclassified against classified for a post-election investigation on their proper operations.

The 18th presidential election in South Korea in 2012, which was a close election with 52% versus 48% for the top two candidates, used the op-scan counters, producing classified (96% of total votes) and unclassified ballots (4% of total votes). Focusing on the top two candidates (hereafter), we noticed a between-candidate relative inequality in the two groups. Candidate 1 from the incumbent party won the vote counts among the classified votes in 161 districts (64%) and won the unclassified vote counts in 208 districts (83%).

We compared vote ratios in each districts using a relative ratio K, defined in section 2.2. Intuitively, if valid ballots have an equal chance of being unclassified by the op-scan counters regardless of candidates, the K-value should be close to 1. It turned out however that the K-value was larger than one in 249 districts (99.2%), and thus candidate 1 always obtained relatively higher votes than candidate 2 from the unclassified group. This apparent favor toward candidate 1 in the unclassified became the main motivation of this study as this unexpected favor for candidate 1 occurred even in electoral districts where candidate 2 received more votes than candidate 1 among the classified. One of the study purposes is to explain the observed difference between two candidates in terms of the relative ratio K.

We set up three study objectives. First, we derive theoretical distributions of the classified and unclassified ballots, and find the theoretical expectation and variance of the proposed relative ratio K (Section 2). Second, we introduce a case study on the 18th presidential election and compare the election results with two previous elections in 2007 and 2002. We then examine the relative ratio K of all districts, construct a national model for the apparent voting pattern, and analyze the impact of the model parameter (K) on winning the election (Section 3).Third, we do a simulation to demonstrate how the national model could be implemented based on the practical process of paper ballot counting by both op-scan counters and counting officials and compare to a systematic op-scan bias (Section 4). We, discuss source of undetectable misdistributions by the op-scan counters, and suggest some bias prevention methods (Section 5). Finally, we conclude with warnings on the op-scan counters in elections.

2. Theoretical Distributions of the Classified and Unclassified Ballots

2.1 Votes sorting process: classified versus unclassified

A picture of the paper ballot with 7 candidates, which was used in the 2012 election, is shown in Figure 1. The ballot seems to be fair, since all candidates have the same area of boxes beside their names (last column in Figure 1) [6]. According to the NEC, the ballot sorting process can be summarized into two stages [11]. In stage 1, the op-scan counters first sort each paper ballot into four categories: candidates 1, 2 and others, and unclassified, which are denoted by P1, M1, Q1, and U, respectively, where Q1 represents votes for the other candidates outside of the top two. When the op-scan counters operate properly, only invalid ballots are expected to be sent to the unclassified. However if the op-scan counters work improperly, classified or unclassified would be mixed with valid and invalid ballots (Figure 2). As the op-scan counters are claimed to be error-free for valid ballots, their misdistribution can happen when they operate by a pre-programmed algorithm as well as random mechanical malfunctions.

In stage 2, the second sorting process is conducted by the counting officials, sorting out the unclassified ballots manually into four categories: candidates 1, 2 and others, and **invalid**, which are denoted by P2, M2, Q2, and U2, respectively (Figure 2). These notations will be used in later sections.

It should be noted that the unclassified and invalid ballots are different, as not all unclassified ballots are invalid. In fact, about 10% of the unclassified turned out to be invalid in the 2012 election, which indicates 90% of them could be sorted back to candidates by counting officials. Also, while misdistribution in the unclassified can be corrected in stage 2, misdistribution in the classified has little chances to be corrected in this voting system.

[Figure 1 here]

[Figure 2 here]

2.2 A proposed measure (K) of between-candidate relative inequality

For each district we let K_C, K_U and K denote three ratios, where K_C is a ratio of the two candidates' received vote counts (or rates), candidate 1/candidate 2, from the classified, K_U is that from the unclassified, and K is the relative ratio of the two ratios. Thus K is a function of the classified and unclassified:

$$\mathbf{K=K_U/\ K_C = (P2/M2)/(P1/M1),}$$

using the notations in Figure 2.

We now focus on **valid ballots only** (excluding invalid ballots), which are unclassified by the op-scan counters such as P2 and M2 in Figure 2. As long as the paper ballot is designed fairly as shown in Figure 1, those valid ballots unclassified should be generated at random, regardless of candidate. We propose the relative ratio K as a measure of between-candidate relative inequality of their valid ballots unclassified by op-scan counters due to unknown reasons.

2.2.1 Theoretical expectation of K: E[K]=1

If valid ballots are unclassified at random, which is fair, the probability of candidate 1 or 2's valid vote to be sent to the unclassified should be the same, noted $\Pr(U_P)= \Pr(U_M)$. Let P=P1 + P2 and M=M1 + M2, where P and M are constants representing the total received votes of the two candidates, respectively. Since each valid vote will be either classified or unclassified independently, we know P1, P2, M1 and M2 all follow binomial distributions as follows:

- P1 ~ B(P, 1-r), where B represents a binomial distribution with a probability r = Pr(U_P).
- P2 ~ B(P, r)
- M1 ~ B(M, 1-r), where r= Pr(U_M)=Pr(U_P).
- M2 ~ B(M, r)

It is known that if X ~ B(n, p), $E\left[\frac{1}{X+a}\right] = \int_0^1 t^{a-1} \cdot P_x(t)dt$, where $0 < p < 1$, E represents expectation, and $P_x(t)$ is the probability generating function. We thus have

$E\left[\frac{1}{X+1}\right] = \int_0^1 t^0 \cdot (q + pt)^n dt = \frac{1-q^{n+1}}{(n+1)p}$, where $q = 1 - p$) [2]. If $n \to \infty, then\ q^n \to 0\ as\ 0 < q < 1$. Thus $E\left[\frac{1}{X+1}\right] = \frac{1}{(n+1)p} \simeq \frac{1}{np}$ as $n \to \infty$.

Also for large X, we see $E\left[\frac{1}{X+1}\right] \approx E\left[\frac{1}{X}\right] \simeq \frac{1}{np}$.

Applying to the 18th presidential election in 2012, we have

$E\left[\frac{1}{P1}\right] = \frac{1}{P(1-r)}$ and $E\left[\frac{1}{M2}\right] = \frac{1}{Mr}$ for large P1 and M2.

Since candidate 1's classified versus unclassified votes are independent from candidate 2's,

$$E[K] = E\left[\frac{P2/M2}{P1/M1}\right] = E\left[\frac{\mathrm{P2}/\mathrm{P1}}{\mathrm{M2}/\mathrm{M1}}\right] = \mathrm{E}\left[\frac{\mathrm{P2}}{\mathrm{P1}} \cdot \frac{\mathrm{M1}}{\mathrm{M2}}\right] = E\left[\frac{P2}{P1}\right] \cdot E\left[\frac{M1}{M2}\right].$$

As P and M are constants, we get

$$E\left[\frac{P2}{P1}\right] = E\left[\frac{P-P1}{P1}\right] = E\left[\frac{P}{P1} - 1\right] = N \cdot E\left[\frac{1}{P1}\right] - 1 = \frac{P}{P\cdot(1-r)} - 1 = \frac{r}{(1-r)},$$

$$E\left[\frac{M1}{M2}\right] = E\left[\frac{M - M2}{M2}\right] = M \cdot E\left[\frac{1}{M2}\right] - 1 = \frac{M}{Mr} - 1 = \frac{1-r}{r},$$

$$E\left[\frac{P2}{P1}\right] \cdot E\left[\frac{M1}{M2}\right] = \frac{r\cdot(1-r)}{(1-r)\cdot r} = 1.$$

Therefore $E[K] = 1$.

Note that $E[K] = 1$ for any electoral district if its size (number of voters) is large enough to use the asymptotic approach. The difference between the theoretical expectation and observed K-value is an indicator of systematic bias in generation of the unclassified votes of each candidate and offers a measure of between-candidate relative inequality with respect to the unclassified ballots. Negligible difference implies negligible bias. Otherwise, one can raise a reasonable doubt on the accurate and fair operation of op-scan counters.

2.2.2 Variance of K depending on electoral district size

In general, there are no closed-form expressions for $E\left[\frac{1}{(X+1)^a}\right]$ for a binomial variable X and a constant a>1, but asymptotic results are available. Cribari-Neto et al. [3] suggested an approximation:

$$E\left[\frac{1}{(X+1)^a}\right] = E\left[\frac{1}{(X)^a}\right] = (np)^{-a} + \left(\frac{a-1}{2p} - \frac{a+1}{2}\right)\frac{\Gamma(\mathrm{a}+1)}{\Gamma(\mathrm{a})}\frac{1}{n^{a+1}p^a} \text{ for large X.}$$

Since P1 ~ B(P, 1-r) and M2 ~ B(M, r), we get

$$E\left[\frac{1}{(P1)^2}\right] \cong \frac{1}{(P(1-r))^2}\left(1 + \frac{1}{P(1-r)} - \frac{3}{P}\right),$$

$$E\left[\frac{1}{(M2)^2}\right] \cong \frac{1}{(Mr)^2}\left(1 + \frac{1}{Mr} - \frac{3}{M}\right).$$

For two independent variables, Y and Z, we have

$$\mathrm{Var(Y)} = \mathrm{E}[\mathrm{Y}^2]-\mathrm{E}[\mathrm{Y}]^2 \text{ and}$$

$$\mathrm{Var(YZ)} = \mathrm{Var(Y)Var(Z)}+\mathrm{Var(Y)E[Z]}^2+\mathrm{Var(Z)E[Y]}^2,$$

where Var(Y) is the variance of Y. Applying the above to the K, we get asymptotically

$$\mathrm{Var(K)} \cong \frac{(P+M)r(1-r)+1}{PM(r(1-r))^2} \cong \frac{(P+M)}{PMr(1-r)},$$

which depends on not only the probability r but also candidates' received votes counts of the electoral district (or electoral district size). Thus the variance of K will be smaller for larger electoral districts, which means the observed K-value should be closer to its expectation. This property of the relative ratio K can be applied to any elections where the op-scan counters are used as primary counting tools.

2.2.3 Lognormal Test for K

Note that K has known mean and variance but unknown probability distribution. To test between-candidate relative inequality in the unclassified group, which is a nonrandom association, a simulation study can be used to fit a parametric distribution. Since K is always positive, we considered a lognormal distribution for the K. In other words, it is a testing if an observed K-value is not different from its expectation based on a normal distribution for log(K) instead of K. For the *i*-th electoral district, K_i has mean 1 and variance V_i, and thus $\log(K_i) \sim N(\mu_i,\ \Sigma_i)$, where $\mu_i = -\frac{1}{2}\log(1+V_i)$ and $\Sigma_i = \log(1+V_i)$. For small V_i, $\mu_i = -\frac{1}{2}V_i$ and $\Sigma_i = V_i$, since $\log(1+V_i) \approx V_i$. This can be applied for the test, as long as the lognormal distribution is a proper fit to K.

For the simulation data were generated based on equal rate of being unclassified for two candidates by three factors: (1) size of electoral district from 1,000 to 100,000 by 1000; (2) candidate 1's received vote rate from 0.2 to 0.8 by 0.1; (3) rate of the unclassified group from 0.02 to 0.15 by 0.01. There were 5,000 runs for each combination, for which Shapiro-Wilk [13] normality test was applied to log(K). The results indicate lognormality of K if the size of electoral district ≥ 10,000, candidate 1's received vote rate ≥ 0.3, and the rate of the unclassified group ≥ 0.03. If these conditions are not satisfied, Fisher's exact test can be used on the vote counts directly.

3. A Case Study: Analysis of the 18th Presidential Election in 2012

3.1 Observed differences between classified and unclassified ballots

All election data were obtained from the NEC of South Korea through official request processes [4,12] There were several datasets related to the 18th presidential election including printed tabulations of the voting data transmitted by the op-scan counters and hand-written tabulations prepared by the counting officials. The former was used in this study, from which information on both classified and unclassified ballots were available.

Some features of the 18th presidential election outcomes seemed unusual when we compared the voting results between the two groups. An interesting feature was the noticeable differences in the two ratios between the two groups (K_C vs K_U). A good example was found in electoral district Guri in Gyeonggi, which had a total of about 110,000 votes with rate of the unclassified 3.3%. While the top two candidates had a very small difference of 0.1%p from the classified (49.9% vs. 49.8%), they showed a quite significant difference of 18.1%p (54.7% vs. 36.6%) from the unclassified, i.e. **K_C=1.00, K_U=1.49 and K=1.49, equivalent to log(K)=0.40**. Applying the variance of the K in section 2.2.2, the standard deviation (SD) of the K is about 0.034, and thus the observed K-value of 1.49 (or log(K) = 0.40) from this electoral district is extremely unlikely from a lognormal distribution with mean -0.0005 and SD 0.034 (p-value < 10^{-12}).

3.2 Observed differences between three recent presidential elections

Since a stronger favor toward candidate 1 in the unclassified (i.e. K>1) was observed consistently from the 18th presidential election, we used previous election data from the 16th and 17th presidential elections in 2002 and 2007, where comparable op-scan counters were used. Unfortunately, data availability was limited to only three districts from these previous elections. Following a public election law established in 1994, election results were no longer documented as the elected president's term finished.

Table 1 shows differences between the two groups for the three specific districts. The three districts in the 16th and 17th elections showed comparable K_C and K_U , and thus the K-values were close to the theoretical expectation 1 (1.02 and 1.04). For example, district G showed the two candidates earned (36.3% vs. 56.5%) and (33.7% vs. 50.3%) from the classified and unclassified, respectively. Similarly, another district Y showed comparable results from the classified (16.9% vs. 59.6%) and unclassified (17.5% vs. 59.4%).

In contrast, the voting outcomes from the 18th election corresponding to the three specific districts showed the K values larger than 1 (1.35-1.44). From all 251 electoral districts of the 18th election the relative ratio (K) overall ranged from 0.97 to 2.17 with mean 1.48 (see Table A1).

The three districts are of relatively moderate or large size, and the SD of the K is as small as 0.03 (subsection 2.2.2). Considering no significant changes in the number of voters and comparable unclassified rates for those three districts among the 16-18th elections, we interpret the increased K-values in the 18th election only as an indicator of between-candidate relative inequality.

[Table 1 near here]

3.3 A National Model

The relative ratio K has been developed for individual electoral districts, which count votes independently at different locations. We are also interested in a national relative ratio K_N over all electoral districts in the 18th election.

3.3.1 National K value for all 251 Electoral Districts of the 18th Election

Since $\mathbf{K=K_U/K_C}$, we first investigated the relationship between $\mathbf{K_U}$=P2/M2 and $\mathbf{K_C}$=P1/M1. Figure 3 shows K_U (y-axis) versus K_C (x-axis) for all 251 districts, where the slope indicates the national relative ratio, K_N (see Appendix Table A1 for the relative ratios).

[**Figure 3 here**]

The mean, median and inverse-variance weighted average of the 251 K-values were 1.48, 1.47, and 1.45 respectively, indicating the K symmetric (Figure 3). Thus the estimate of K_N is 1.5. In equation, the national model is simplified as follows:

$$\mathbf{P2/M2= 1.5* (P1/M1).} \qquad \mathbf{(1)}$$

To illustrate how the model works, we assume that the vote ratio in the classified is one (P1/M1=1 or simply K_C=1), which means no difference between the two candidates. In such a situation, multiplying by 1.5, P2/M2 becomes 1.5, i.e. 3/2, and thus 60% (=3/5) of the votes goes for candidate 1, whereas 40% (=2/5) for candidate 2 in the unclassified. Therefore, the between-candidate difference has increased from 0%p (classified) to 20%p (unclassified). In fact, a very similar voting result was observed in district Guri City with differences of 0.1%p versus 18%p from the classified and unclassified, respectively.

3.3.2 Impact of the national model parameter (K) on winning an election

As model (1) assigns more votes to candidate 1 than candidate 2 in the unclassified proportionally to the classified regardless of who wins in the classified, we face a critical question: ***what is the impact of the model parameter on the election outcome?***

To answer the question, we let **L=P1/M1** and P2/M2=K*(P1/M1), where K ≥1 and L >0. Then P2/M2=K*L, and thus P1=L*M1 and P2=K*L*M2. For simplification without loss of generality, we take candidate 1's perspective of winning the election. If candidate 1 gets more votes than candidate 2, it clearly means that (P1+P2) > (M1+M2), and thus we set up candidate 1's winning condition as follows:

$$\mathbf{(L-1)*M1+(K*L-1)*M2>0.} \qquad \mathbf{(2)}$$

The only non-trivial case for candidate 1 winning via the unclassified ballots is as follows: If $L<1$ (i.e. $P1<M1$), then $(K*L-1)*M2 > (1-L)*M1$. Since $M2>0$, we have $(K*L-1) > (1-L)* (M1/M2)$. To simplify this inequality, we set **x=M1/M2**, where x is a ratio of two votes for candidate 2 between the classified (M1) and the unclassified (M2). Then equation (2), which is the candidate 1's winning condition, becomes a function of x and K such that $L > (1+x)/(K+x)$. This is the interesting case to be discussed further.

Candidate 1's winning depends on two variables, x and K, even if $K >1$. We fix $K = 1.5$ and examine the effect of the model parameter ($K=1.5$), varying the values of x. Since the votes for candidate 2 from the classified are larger than the unclassified (i.e. $M1>M2$ from the 18th election), the minimum value of x is 1 ($x\geq1$).

As demonstrated in Figure 4, model (1) increases candidate 1's winning as the number of the unclassified ballots goes up. Since the unclassified are to be generated by the op-scan counters, the model parameter K value is linked to the op-scanner's operation. The impact of the model parameter (K) on candidate 1's winning would be maximized when candidate 1 lost in the classified ballots.

In the 18th election, x was observed close to 35 and $L=0.99$, which indicates candidate 1 could win the election with less votes from the classified (at least 99% of candidate 2) but more votes from the unclassified. This could be achieved by setting $K=1.5$ and thus elucidates the meaning of the model parameter 1.5 in the 18th presidential election.

[Figure 4 here]

4 Simulation and Results

Even though the national model (1) can explain the actual voting outcomes quite well, there still remains a question on how it can be implemented in the actual votes. We performed simulations with two scenarios to demonstrate potential unintentional systematic bias and intentional systematic manipulation of the op-scan counters. The first scenario involved a systematic bias in unclassified rate, where valid votes for candidate 1 are more likely than others to be unclassified, but within a pre-set tolerance that may not be detectable in smaller quality assurance test prior to a full election. For the second scenario we considered two facts: (a) the op-scan counters are electronic computing devices, thus programmable; (b) only the unclassified were manually sorted later by the counting officials [5,11].

According to the NEC [10], the op-scan counter is accurate in ballot counting as it is claimed to be an error-free sorting device. Thus, for the second scenario, we assumed that all op-scan counters operate properly following how they are set-up or programmed. This simulation is based on a scheme with two types of misdistribution generated by the op-scanners. First, valid ballots for candidate 1 or 2 are sent to the unclassified (**first misdistribution**), which was actually observed in the 18th election. This first misclassification would result in a non-negligible shortfall of each candidate in the classified, and in this simulation the invalid votes only are used to fill in the shortfalls (**second misdistribution**). Both misdistributions are clearly the source of incorrect vote counting. In this section we explain briefly simulation assumptions, process, and results.

4.1 Scenario 1: unintentional systematic bias

A potential unintentional systematic bias ($\beta>0$), which can explain the national model (1), was set up using the notations in subsection 2.2.2: P2 ~ B(P, r+β), whereas M2 ~ B(M, r). Here r and β are fixed, so that the overall unclassification rate is 3.7% as in the real election data. It is required 1+ β/r be approximately 1.5 to obtain K's within the range of the ones observed in each district. We set r=0.03 and β=0.0145 for this systematic bias scenario.

4.2 Scenario 2: intentional systematic manipulation

Assumptions required for scenario2 include: (1) The vote rate for the other candidates except for candidates 1 and 2 is very small and negligible (0.37% of the total votes); (2) The percentage of the unclassified (R) is small, varying over districts (nationally, the classified vs. unclassified is overall 96:4); (3) The percentage of invalid votes (R2) is 10% of the unclassified (nationally and fixed for all 251 districts); (4) The vote rate for the other candidates in the unclassified is also very small and negligible (1.3% of the unclassified); (5) Correction of the first misdistribution is properly done by the counting officials following fair rules set up by the NEC; (6) Correction of the second misdistribution is not done by the counting officials.

This simulation has two stages following the ballot sorting process in Figure 1. In the first stage, virtual ballots are created to be close to the actual voting results. Prior information required here is: total number of votes (classified + unclassified + invalid vote), received vote rate of candidates 1 and 2, respectively, and percentage of the unclassified to the total votes. The virtual ballots can be created using multinomial distribution. Out of 1000 trials, the best will be kept, which is the closest to the actual election data over all districts. In the second stage, the op-scan counters sort out the virtual ballots in accordance with the given conditional probabilities (Appendix B). Conditional Bernoulli (a) and multinomial (b) distributions are used with pre-assigned probabilities for this classification as follows:

a) If a virtual ballot is for candidate *j*, it will be sent to either candidate *j* or unclassified, for *j*=1,2 or other.

b) If a virtual ballot is an invalid vote, it will be sent to candidates 1, 2 or unclassified.

[Figure 5 here]

4.3 Results

The first simulation scenario involving a systematic bias, while providing similar K values, does not match the district level variability in the unclassification rates. The motivation is that r+ β is within some tolerance limit so that the candidate level examination would be omitted in the calibration of the op-scan and the bias would go unnoticed before the election. The simulation results based on the systematic machine bias model do little support plausibility in the 18th election (Table 2), mainly because unclassified rates vary among districts and exhibit a non-uniform but unidirectional bias in favor of candidate 1.

In comparison, the second simulation shows the simulated votes are fairly close to the actual votes nationwide (Table 2). Considering different population size between large and small electoral districts, we evaluate simulation results for each of 251 electoral districts in received vote rates (%) rather than received vote counts (see Appendix Table C1 for all 251 districts). The comparable results over all districts imply that the simulation reasonably reflects the actual votes and thus can explain a way of model implementation. The simulation results turned out that 97% of 251 electoral districts showed very good predicted vote rates within ±5% of margin of acceptance (Appendix Table C2). This indicates the proposed scenario 2 can describe a plausible way by which the op-scan counters may have been operated in the 18th presidential election.

[Table 2 here]

5 Discussions

5.1 Interpretation of simulation results

The relative ratio K was proposed as the measure for between-candidate relative inequality when classified and unclassified votes were generated by op-scan counters. Since the op-scan counters are programmable devices, we compared simulations (scenario 2) involving bias caused by programming the op-scan counters and by systematic but hard to detect calibration problem (scenario 1). Note that there could be multiple ways for applying model (1) to the votes counting process, and thus simulation scenario may be **not unique**. The proposed simulation carries out model (1) using invalid ballots only, minimizing the impact of the model on the election winner. If valid votes for candidate 2 were sent to candidate 1 or vice versa, the impact of the model (1) could affect the election up to changing its outcome, but it would be more easily detectable. Shifting invalid votes, which should be few, would constitute a small hedging of one candidate's chances in a close election.

The most important finding from the simulations is that the op-scan counters can generate serious misclassifications that are better explained by a pre-programmed algorithm, than by systematic unintentional, bias or random mechanical malfunctions. Since all misclassifications in the simulation took place in the sorting process by the op-scan counters, the proposed simulation design can be interpreted as a warning for more openly and thoroughly tested use of op-scan counters in elections with post-election auditing results from the machines.

5.2 Programmable but undetectable misclassifications by op-scan counters

Most electoral fraud or manipulations have been conducted locally [8,9]. We claim in this study a nationwide potential manipulation in a presidential election using the op-scan counters. As programming the op-scan counters can be done behind scenes, this process is unlikely to be found or detected by the election observers, whose task is to ensure fair votes counting by all means. In spite of the benefits of using op-scan counters, a thorough evaluation on whether or not to continue to use these op-scan counters is essential to eliminate potential rigged vote counting.

Solutions for detecting election fraud have been suggested by auditing a well curated paper trail against the electronic results [14] or auditing a random sample of the ballot boxes

[1,14,15]. Vote tabulation audits can serve process monitoring, quality improvement, fraud deterrence, and bolstering public confidence [14]. Serious errors can go undetected if results are not audited effectively [1,7,14,15].

Statistical methods can be more effective than auditing in detection of between-candidate relative inequality when the op-scan counters are used. We proposed a measure, the relative ratio K, to detect between-candidate relative inequality. If the K-value is not close to its expectation 1 for any electoral district of sufficient size, it may indicate that valid ballots unclassified by the op-scan counters were attributed unevenly to candidates and thus the election results become disputable, demanding further investigation. The proposed relative ratio K can be used for both targeted and extensive post-electoral auditing according to how localized or widespread observed deviations are from a fair electoral model. In the 2012 election the 251 K-values were around 1.5 much larger than their expected value of 1 across nation. Potential causes of this remarkable election outcome can be either op-scan counter related manipulations or unknown equipment related bias. The three specific electoral districts, of which the K-values close to 1 in the 16^{th} and 17^{th} elections became much larger than 1 in the 18^{th} election, seem to support the former rather than the latter.

The K-value can be examined for some electoral districts individually for local, regional, or national analysis. We also proposed a national model to detect systemic issues rather than local outliers. Applying both methods, we are able to detect between-candidate relative inequality in election at local, regional or national levels.

We provide the post-election data and simulation codes as supplemental materials (Appendices A to D) to promote public interest in election votes counting system, to have more extensive data analysis from fellow researchers, and ultimately to reach higher level of public awareness of invaluable fair and accurate votes counting.

6. Conclusions

The strength of this study is being able to demonstrate a potential serious loophole in using the op-scan counters, which can be **error-free but not manipulation-free**. The proposed measure of between-candidate relative inequality (K) and national model over all electoral districts could contribute to securing and promoting of accurate and fair votes counting of upcoming worldwide elections, where the op-scan counters are to be used as the primary main tools for votes counting. Future development of the measure include further theoretical considerations and more sophisticated modeling to take into account other potential sources of bias or misdistribution of votes by the op-scan counters.

References

[1] A.W. Appel, M. Ginsburg, H. Hursti, B.W. Kernighan, C.D. Richards, G. Tan, and P.Venetis, *The New Jersey voting-machine lawsuit and the AVC advantage DRE voting machine*, EVT/WOTE'09, 2009.

[2] M.T. Chao, and W.E. Strawderman, *Negative moments of positive random variables*, J. Amer. Statist. Soc. 67 (1972), pp. 429-431.

[3] F. Cribari-Neto, N.L. Garcia, and K.L.P. Vasconcellos, *A note on inverse moments of binomial variates*, Braz. Rev. Econom. 20 (2000), pp. 269-277.

[4] B. Chung, *Post Presidential Election Data in 2002 & 2007 (NEC disclosure)*, 2016.

[5] B. Chung, *The Accusation Against 18th Presidential Election Fraud*, Baweesol, 2013.

[6] J. Elklit and A. Reynolds, *A framework for the systematic study of election quality*, Democratization 12 (2005), pp.147-162.

[7] A.J. Feldman, A. Halderman, and E.W. Felten, *Security Analysis of the Diebold AccuVote-TS Voting Machine*, USENIX/ACCURATE EVT'07, 2007.

[8] P. Klimek, Y. Yegorov, R.Hanel, and S. Thurner, *Statistical detection of systematic election irregularities*, Proc. Natl. Acad. Sci. USA. 109 (2012), pp. 16469-16473.

[9] D. Kobak, S. Shpilkin, and M.S. Pshenichnikov, *Integer percentages as electoral falsification fingerprints*, Ann. Appl. Stat. 10 (2016), pp. 54-73.

[10] National Election Commission (NEC) of South Korea, *Op-scan counter is accurate and fast*, (2014), Available at http://blog.nec.go.kr/220008433572 .

[11] National Election Commission (NEC) of South Korea, *Handbook of Public Election Process*, 34-9760878-100012-14, 2014.

[12] Project Boo, *Post Presidential Election Data in 2012 (NEC disclosure)*, Feb.28 2017; data available at https://drive.google.com/open?id=0Bx0QzBOFfX95MkI3MmcwUHhCMVk

[13] S.S. Shapiro and M.B.Wilk, *An analysis of variance test for normality (complete samples)*, Biometrika 52 (1965), pp. 591–611.

[14] P.B. Stark, *Risk-limiting vote-tabulation audits: The importance of cluster size*, Chance 23 (2010), pp. 9-12.

[15] B. Wofford, *How to hack an election in 7 minutes*, Politico Magazine, 2016.