WHY DID THE PRE-ELECTION POLLS IN THE SOUTH KOREAN LOCAL ELECTIONS GO ALL WRONG? ASSESSING THE SOURCES OF ERRORS USING A DUAL-FRAME LANDLINE/CELL PHONE POST-ELECTION SURVEY

Survey & Health Policy Research Center
Technical Report

March 2018

dongguk UNIVERSITY SHPRC

WHY DID THE PRE-ELECTION POLLS IN THE SOUTH KOREAN LOCAL ELECTIONS GO ALL WRONG? ASSESSING THE SOURCES OF ERRORS USING A DUAL-FRAME LANDLINE/CELL PHONE POST-ELECTION SURVEY

Sun-Woong Kim, **Michael W. Traugott**, **So-Hyung Park**, **Sang-Kyung Lee**
Dongguk University, University of Michigan, Dongguk University, Hyundae Research Institute

Abstract

Election polls provide information about where a race stands, and their estimates of candidate preference may influence the public prior to the election. If pre-election polls err in estimating vote shares, they can mislead the public and affect their attitudes and behavior because of their reliance on polls. In the 2010 South Korean local elections, all the pre-election polls misestimated the outcome. After the elections, many pollsters, newspapers, and political experts ascribed poll failures to issues like the spiral of silence or out-of-date methodology, but they offered no evidence to support any of these explanations. We present a scientific evaluation of the sources of errors in the polls, based on the results obtained from a dual-frame landline/cell phone post-election survey and suggest several ways to improve pre-election poll accuracy.

Keywords: Noncoverage, Sampling error, Nonresponse, Landline and cell RDD, Sample weights, Self-reported voters

Address correspondence to Sun-Woong Kim, Department of Statistics, Dongguk University, 26 3Ga, Pil-Dong, Jung-Gu, Seoul, South Korea, 100-715, e-mail: sunwk@dongguk.edu

Introduction

Although the number of election polls around the world has been increasing with the development of new survey methods, pollsters do not estimate all election outcomes with equal success. While the quality of estimation at the national level is quite good and differences between the estimates and election outcomes seem to be declining over time in several countries, many pollsters and researchers often suffer from poll failures when estimating election outcomes at the sub-national level. Recent examples include polling for the 2008 primary elections in the United States, the 2010 gubernatorial elections in Mexico, the 2007 and 2008 provincial elections in Canada, and the 2006 and 2008 general elections in Italy (Traugott & Wlezien, 2009; Durand, 2011; Moreno, Aguilar, & Romero, 2011; Sala & Fumagalli, 2010; Durand, Foucault, Goyder, & Deslauriers, 2010). The standard questions are whether there were methodological problems or socio-political circumstances in each country that contributed to the estimation errors.

From a methodological perspective, pre-election polling is a process that involves several steps in order to produce an estimate, and problems or errors can arise at any one of them. There are design issues involved in the decision about when and how to field a study. Research shows that estimates are more accurate when the fieldwork is closer to Election Day (Crespi, 1988). Most studies are now conducted on the telephone, but mixed mode designs have certain advantages (Atkeson et al., 2011). With the rapid adoption of cell phones, issues of how to design samples to incorporate both landline and cell phone users create new problems because of differences in the demographics and hence the political characteristics of those who use each (Keeter et al., 2007). Under any circumstance, not every respondent votes; pollsters must assess the likelihood that each respondent will vote through a series of questions (Bolstein, 1991; Freedman and Goldstein, 1996). Analysts combine the responses into a likelihood measure that is often considered proprietary and therefore not made publicly available. Even the question wording itself can affect the level of overreporting of intended voting behavior, a well-known phenomenon observed in pre-election polls in every country (Karp and Brockington, 2005).

Aspects of the fieldwork can also create problems, starting with the response rate and the associated non-random loss of potential voters. Media political polls often involve only a limited field period that constrains the number of callbacks that are feasible (Traugott, Groves, & Lepkowski, 1987). The simple act of being interviewed can create conditions for socially desirable responses, as individuals do not want to tell a stranger that they do not intend to vote, they prefer the underdog in the election, or they will not vote for a minority or female candidate (Hopkins, 2009).

In the research reported here, we focus on methodological issues in the design of pre-election polls in South Korea because there are a number of identifiable problems in the current survey work done there. We focus primarily on the sampling frame and the use of weighting and their effects on estimation. This does not mean that these are the only explanations for estimation error, although our results suggest that they could account for a significant reduction in it. At the same time, we collected our data through a post-election study conducted a few months after the election. Many academics conduct post-election studies of voting behavior, mainly to explain voting patterns but not to investigate methodological issues. A proper evaluation of the methodological determinants of estimation error would involve testing *a priori* hypotheses evaluated through experimental manipulations in the design, questionnaire, and fieldwork for pre-election polls. We believe our findings inform the design of such future studies.

The South Korean Electoral System and Media Environment

South Korea has a multi-party system with candidates appearing on the ballot with their own name and party affiliation. There are several parties with local strength, but only a few have significant national drawing power. For example, in the 2007 presidential election, three candidates – from the Grand National Party (GNP), the United New Democratic Party (UNDP), and an Independent candidate – received 89.9% of the vote. In the 2012 National Assembly elections, candidates from three parties – the Saenuri Party (SP), the Democratic United Party (DUP), and the United Progressive Party (UPP) – received 89.6% of the vote. In the 2010 local elections, the most successful candidates were from the GNP and the Democratic Party (DP), winning almost all of the seats between them.

In South Korea, news organizations often sponsor polls to contribute to their news coverage, especially during election campaigns. Sometimes they do this on their own, and occasionally they collaborate with other news organizations or survey firms in sponsoring polling. It is important to note that there is a blackout period for reporting poll results in the week before a Korean election, and all of the news organizations that sponsor polls abide by it. The time series of polling data that precedes any election typically contains results from multiple news organizations and polling firms, creating an issue of potential "house effects" as a source of reported differences in their estimates (Smith, 1978 and 1982).

Reporting styles associated with the publication of poll results and the horse race nature of campaign coverage can also affect citizens' expectations of the outcome of an election. For example, before the election, *Dong-A Ilbo*, one of the three major conservative papers, reported "In Seoul, Oh (GNP candidate) would beat Han (DP candidate) by 20.8% percent." (Dong-AIlbo, 2010a). Also, the Korean Election Pool (KEP) sponsored by the three major broadcast networks (KBS, SMC, SBS) reported "Showing the GNP Ahead Comfortably in the Big Three Regions (i.e., Seoul metropolitan area including Seoul, Incheon, and Gyeonggi)" (ChosunIlbo, 2010a). After the election, both the Korean and foreign press reported on the poor performance of the polls and the explanations offered by some pollsters for their difficulties. For example, in summarizing what happened, the *JoongangIlbo* reported "In many regions, GNP candidates were forecast to take sweeping wins, especially in the Seoul metropolitan area. The election results, however, gave close - or in some cases more comfortable - victories to the DP."(2010).

The Problems of the Pre-election Polls in the 2010 Local Elections

In the case of the 2010 gubernatorial and mayoral elections in South Korea, the pre-election polls produced estimates at the provincial and mayoral levels. In order to understand the dynamics of polling in the campaign, it is useful to consider the contest for the mayor of Seoul, the largest city in Korea. Candidates seeking office had to register by May 14, and campaigning officially started on May 20. Election Day was June 2, two weeks later, and the blackout on publishing poll results started on May 25. The data presented in Figure 1 show the poll estimates over time for this race, reported to the .1% as the Korean pollsters typically do. The *Seoul Shinmun* fielded the first poll on May 6 and reported the results on May 10, before the candidates formally declared; there were six poll estimates of the race before it legally began. In all 14 separate estimates were reported in the media from polls conducted by seven different sponsors across this five-week period. The newspapers reported their poll results separately, while the television networks sometimes reported them on their own or together.

The overriding impression from the graph is the stability of the estimates and the lack of differences between organizations. Across poll dates, the average reported support for the GNP was 49.6% while the average support for the DP was 32.3%. The GNP lead ranged from 11.9 to 22.8 percentage points. There were four televised debates between the mayoral candidates in Seoul, three held during the polling period (May 17, 18, and 19) and one during the polling blackout period (May 28). Some of the polls conducted just after the third debate showed a slight widening of the lead for the GNP candidate as the support for the DP candidate dropped slightly. There were no polls published after the fourth debate by law, and the press reporting of the race did not suggest that the candidates' performance in that debate affected voter support for any of them.

When the votes were counted on Election Day, the GNP candidate received 47.4% of the vote in Seoul while the DP candidate received 46.8%, a very close race decided by less than 27,000 votes out of a total of slightly more than 4.4 million casts. Estimation of the outcome was not a problem for the exit poll sponsored by the three major broadcasters, which showed 47.4% for the GNP and 47.2% for the DP. This raises the question of why all of the pre-election polls were so far off in their estimates of the outcome, specifically in the level of support for the DP in Seoul.

The poll performance was similar but worse in Incheon, the third-largest city in South Korea. As shown in Figure 2, there were also 14 pre-election polls reported in the media. The first was in the field on April 26; there were six completed before the formal start of the campaign and eight during the campaign period itself. Again, the poll estimates over time look relatively stable, with the average support for the GNP at 42.2% and at 32.6%for the DP. The average lead in this period was 9.6 percentage points for the GNP. In Incheon, there was only one debate between the candidates, held on May 25, just as the blackout period began, so there was no polling after it. The final vote tabulation showed that the DP candidate received a majority (52.7%) of the votes. With a winning margin of 8.3 percentage points, the pre-election polls in Incheon consistently underestimated his support by approximately 20 percentage points.

While the previous discussion emphasized the estimation problems in Seoul and Incheon, the same type of errors appeared in the other districts where there were mayoral or gubernatorial elections. These districts ranged in size from less than half a million voters to almost nine million. Turnout was relatively high compared to recent elections, ranging from 45.9% to 65.1% across these districts. In the races, the candidates from parties other than the GNP fared better, winning 9 of the 14 offices at stake. In most cases, the differences in the winning candidate's estimates from the actual result ranged from +3.0 to -26.7 percentage points for the winning candidate and +2.6 to -20.8 percentage points for the second-place finisher, both by more than sampling error alone would suggest.

Right after the elections, many newspapers and broadcasters reported that the major cause of the poll failures was a "spiral of silence effect" (Noelle-Neuman, 1974) resulting in an unmeasured "silent majority" who suspected that the party other than their preferred one would win the election. Instead of giving their true preference, journalists suggested some respondents expressed support for the leading party and its candidate or indicated that they had not made up their minds. But none of the journalists or news organizations provided any empirical evidence for a spiral of silence effect. Their poll results consistently suggested significant wins for the GNP and underestimated support for the DP across time and jurisdiction. This occurred whether the DP won or lost a particular election and suggests the possibility of systematic error that could come from a number of sources. However, the pollsters neither provided copies of their data nor disclosed detailed methodological information about how they conducted their polls.

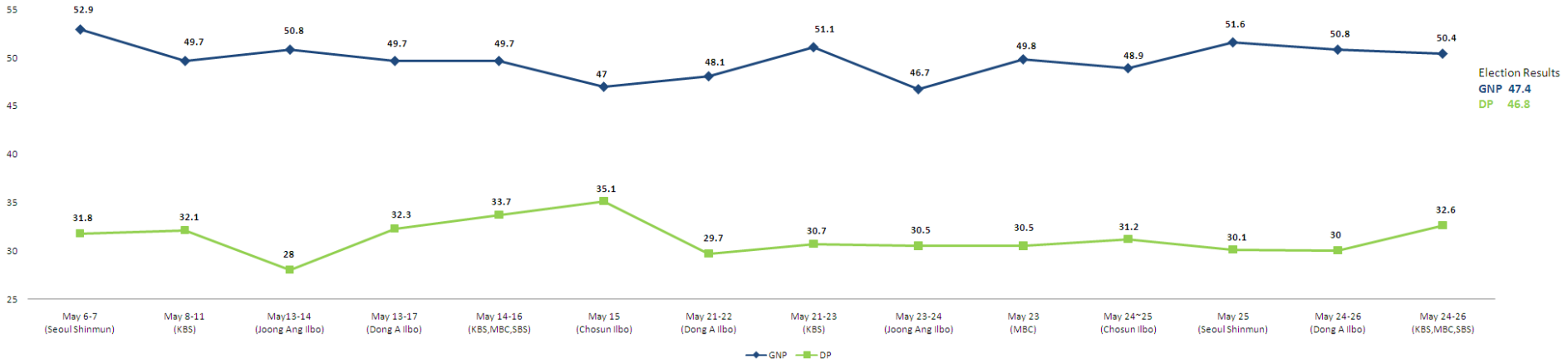


Figure1. Reported Poll Results Preceding the 2010 Seoul Mayor's Race

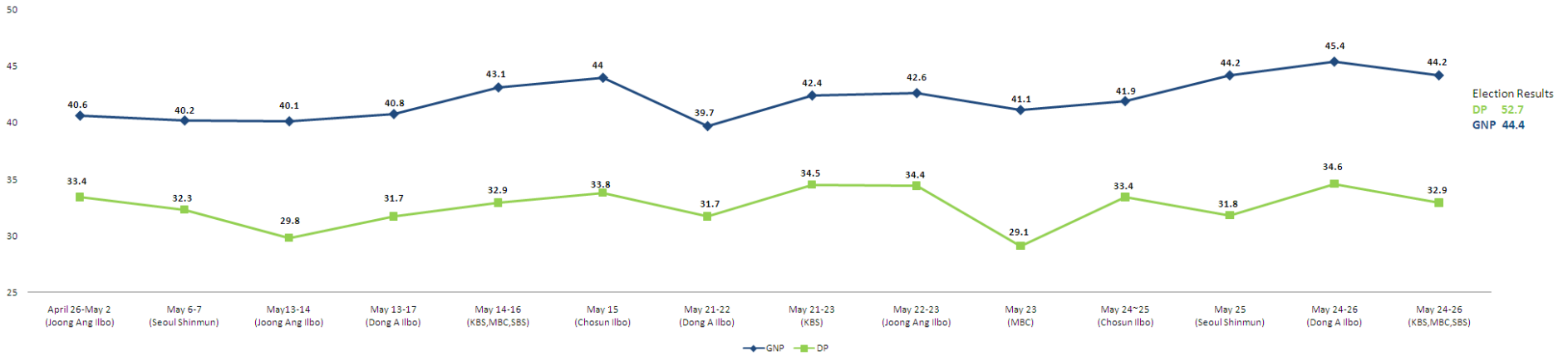


Figure2. Reported Poll Results Preceding the 2010 Incheon Mayor's Race

Research Design

As a follow-up to this controversy, we designed a dual-frame landline/cell phone post-election survey to investigate some of the possible explanations for these estimation problems. We focused on sampling and weighting issues because of some known attributes of polling procedures in Korea, although there may be other explanations as well. It is always difficult to conduct *post hoc* analyses of the possible causes of estimation errors; nevertheless, we believe that our results are useful and point to possible changes in survey methods that could improve estimation in future pre-election polling procedures in Korean elections and elsewhere.

South Korean Election Polling Methodology

Prior to designing our own study, we reviewed the available information about the methodology used in these local election polls. The most striking feature is that the sampling in most polls is based upon a frame of listed landline telephones excluding unlisted numbers or cell phones. Accordingly, there were no interviews with individuals who are cell phone only (CPO) or primarily cell phone (PCP) users, people who have access to both but rely upon their cell phone for communication. This omission of those with unlisted numbers or cell phones is cited in many other countries as a source of problems in pre-election polls or other surveys (Gabler & Häder, 2002; Nicolaas & Lynn, 2002; Keeter, 2006; Keeter et al., 2007; Kalsbeek & Agans, 2008). In South Korea, the percentage of households with both landlines and cell phones was estimated at 75.1%, while that of cell phone only households was 19.5% in 2009 (Choi, Kim, & Cho, 2011; Choi, Kim, Cho, & Couper, 2011). The percentage of unlisted landline numbers was about 50% in 2010 (Kim, Park, & Hong, 2011).

Some Korean pollsters used RDD sampling techniques for telephone number selection, but all of them employed quota sampling according to the area, sex, and age groups (i.e., 20-29, 30-39, 40-49, etc.) for the selection of individual respondents. This methodology provides neither an unbiased estimate nor a valid variance estimate, since the sample is non-probabilistic (Cochran, 1977; Lohr, 1999). None of their field periods were over five days, allowing for only a limited number of callbacks. All of the polls used simple estimation methods that did not depend on weighting or any other adjustments at the household and person level. The sample size per pre-election poll in these races ranged from 500 to 1,000, somewhat smaller than typically used in national public opinion studies in South Korea that range from 1,000 to 1,500. However, these sample sizes do not suggest the possibility of unusually large random errors.

Based on this information about the pre-election poll methodology used in 2010, we designed our study to investigate a variety of sources of error. These included *noncoverage*, due to missing unlisted landline phone numbers or cell-only populations; *nonprobability sampling*, that is, the use of quotas for respondent selection; and *nonresponse*, due to the small number of call attempts per sampled telephone number. Groves (1989) in particular classified errors attributed to these sources as errors of nonobservation. We can also add to the other sources such as *timing*; *question wording*; *response bias*, due to social desirability; and *turnout estimation methodology*.

Survey Design for the Post-Election Study

There are four key elements of the design for the post-election study. First, in order to examine the differences in and impact of listed landline, landline RDD, cell RDD, and dual RDD frames, respectively, we conducted a landline/cell phone RDD survey, obtaining separate samples from each frame. Second, the respondents in our study were randomly selected within a household rather than by quota. Third, we allowed for enough time in the field to make a large number of call attempts. Finally, we used a new weighting strategy for the landline/cell phone survey design (Park *et al*., 2011). The details on design and weighting procedures are given below.

We adopted the following survey design for reducing non-coverage, sampling error, and non-response, respectively. First, we used list-assisted RDD sampling according to Kim *et al*. (2012) for a landline sample combined with an RDD design for cell phones to reduce noncoverage. Second, we randomly selected a respondent among household members using the same phone number, regardless of the type of phone on which initial contact was made (landline or cell). Third, we used a minimum of 10 calls per phone number during weekdays and weekends across a 47-day field period from November 1 to December 17 in 2010 in order to reduce non-response. Fourth, we defined three geopolitical strata, that is, a "GNP stratum" that included the six areas they won, a "DP stratum" that included the seven areas they won, and an "Others stratum" based upon the three areas that independents and minor party candidates won. We ended up with a total sample size of 1,508 (899 landline RDD frame respondents and 609 cell RDD frame respondents). The landline RDD respondents include 465 (51.7%) who would have appeared on a listed landline frame since it is included in the landline RDD frame.

We assume that both landline and cell numbers are either for the household (all household eligibles can be reached at that number), b) shared (more than one eligible person, but not all eligibles, can be reached at that number), or c) personal numbers. Therefore we asked a few questions to identify such status from the informant for each phone number selected from the landline RDD or cell RDD frame. If the phone number was for the household or a shared number, one eligible person using the phone number was randomly chosen and asked to provide the information on other phone numbers that could have been used to reach them (e.g., How many landline or cell phone numbers do you use, and for each phone number how many other people use that phone number?). If the phone number was for a single person, that person was interviewed. This method was used to avoid complicated estimation procedures used to combine results from the landline and cell frames when there was overlap (e.g., an individual could have been selected in either the landline or cell frames). As a result of these procedures, the following simple sample weights for the dual-frame design for landline only persons, cell-only persons, and people with access to both a landline and a cell were used:

1) Landline only person: $1/\pi_{\mathrm{Li}}$
2) Cell only person: $1/\pi_{\mathrm{Ci}}$
3) Landline and cell person: $1/(\pi_{\mathrm{Li}} + \pi_{\mathrm{Ci}} - \pi_{\mathrm{Li}}\pi_{\mathrm{Ci}})$

where $\pi_{Li} = \sum_{j=1}^{\alpha_{Li}} \frac{p_{Li}}{\beta_{Lj}}$

$$\pi_{Ci} = \sum_{j=1}^{\alpha_{Ci}} \frac{p_{Ci}}{\beta_{Cj}}$$

α_{Li}: Number of landlines that could reach respondent i

α_{Ci}: Number of cell phones that could reach respondent i

β_{Li}: Number of adults who use j-th landline with respondent i

β_{Ci}: Number of adults who use j-th cell phone with respondent i

p_{Li}: Selection probabilities of landline RDD numbers

p_{Ci}: Selection probabilities of cell phone RDD numbers

Survey Measures

The main survey items consisted of the following questions. Interest in politics (asked of the entire sample) was determined by asking the following question with the associated response categories: "In general, how interested in politics would you say you are (Very interested, Somewhat interested, Moderately interested, Not too interested, Not interested at all)?"

Self-reported registration and voting (asked of the entire sample) were determined by asking the following questions with the associated response categories: "Were you registered as a voter in the June 2 local election for mayor or governor this year? (Yes, No)"; "Did you vote in the local election? (Yes, No)"

Voting for the winner (asked only of self-reported voters) was determined by asking the following question with the associated response categories: "Did the candidate you voted for win the election? (Yes, No)" Talking with anyone about the elections before voting (asked only of self-reported voters) was determined by asking the following question with the associated response categories: "Did you ever hold a conversation with anyone about the election before you voted? (Yes, No)" The survey also contained questions measuring personal demographics such as age, sex, and education, which were asked of the entire sample.

We acknowledge that, like pre-election polls, our post-election survey could be subject to a number of sources of errors, including the amount of time elapsed since the election, its possible effect on recall of voting behavior, coverage, sampling, non-response, question wording, and response bias due to social desirability for respondents who knew the outcome of the election. But the estimation problems of the pre-election polls only became clear after the votes were counted, and there was no alternative to a post-election study to assess possible causes.

Results

According to the Korean National Election Commission (KNEC), the turnout in the 2010 local elections was 54.5%. With reasonable response rates,[1] the self-reported turnout among survey respondents was estimated at 67.0%, higher than the official rate because the sample of those interviewed is typically composed of individuals with higher socioeconomic

characteristics than the population as well as because of social desirability pressures to respond as a "good citizen."

Table1.
Weighted Sample Demographics of Self-Reported Voters by Frame at the National Level

	Listed Landline	Landline RDD	Cell RDD	Dual RDD	2010 Census*	KNEC**
Gender						
Male	41.1%	35.7%	56.8%	47.6%	48.9%	49.5%
Female	58.9%	64.3%	43.2%	52.4%	51.1%	50.5%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
N	335	624	400	1024		
Age						
19-29	7.4%	8.0%	17.0%	14.1%	17.9%	14.9%
30-39	6.4%	14.0%	20.6%	19.5%	21.2%	17.7%
40-49	22.8%	28.8%	25.1%	25.9%	22.3%	22.5%
50-59	23.9%	19.2%	16.6%	17.7%	17.9%	20.2%
60 or over	39.4%	30.0%	20.7%	22.8%	20.7%	24.7%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
N	335	624	400	1024		

Note. The "Listed Landline" frame and unlisted landline numbers are included in the "Landline RDD" frame. The respondents in "Listed Landline" and those with unlisted landlines are in both "Landline RDD" and "Dual RDD." Those in "Cell RDD" are in "Dual RDD," not in "Listed Landline" or "Landline RDD," since they were selected from the cell phone frame.
* The 2010 census results
**Election study result conducted by the Korean National Election Commission (KNEC) after the local elections with a sample size of 4,033,027 (10.4% of whole actual voters)

Data are presented in Table 1 showing the demographic characteristics of the self-reported voters in the different subsamples in relation to two external measures: the 2010 Korean census conducted by Statistics Korea and a major post-election assessment of voters conducted by the KNEC. Respondents from the cell phone RDD sample are more likely to be male than the respondents from the two different landline samples. When the three subsamples are combined appropriately with the new weighting system, the resulting sample is a much better reflection of the proportion of males in the population according to both the census and the KNEC. The same is true for the age distributions in the samples. The cell phone sample is much younger than the two landline samples; in particular, the sample from the listed landline frame is much older than the landline RDD or cell phone RDD subsamples. This analysis was replicated for respondents who said they voted for the winner in their district, and the same patterns appeared. The cell phone RDD subsample of self-reported voters for the winner was much more male than the two landline samples, and it was much younger as well.

Table2.
Attitudinal and Behavioral Characteristics of the Sample of Self-reported Voters by Frame at the National Level

	Listed Landline	Landline RDD	Cell RDD	Dual RDD
Interest in Politics				
Very	6.3%	5.1%	6.7%	5.4%
Somewhat	13.8%	17.4%	20.8%	20.3%
Moderately	28.9%	35.5%	35.5%	33.8%
Not too	34.5%	29.7%	24.3%	27.1%
Not at all	16.5%	12.3%	12.7%	13.4%
Total	100.0%	100.0%	100.0%	100.0%
N	335	624	400	1024
Talk with Anyone				
Yes	59.4%	59.6%	65.0%	61.0%
No	40.6%	40.4%	35.0%	39.0%
Total	100.0%	100.0%	100.0%	100.0%
N	335	624	400	1024

Note. See the note in Table 1.

Table3.
Self-Reported Vote for the Winner by Frame in the Three Geopolitical Strata

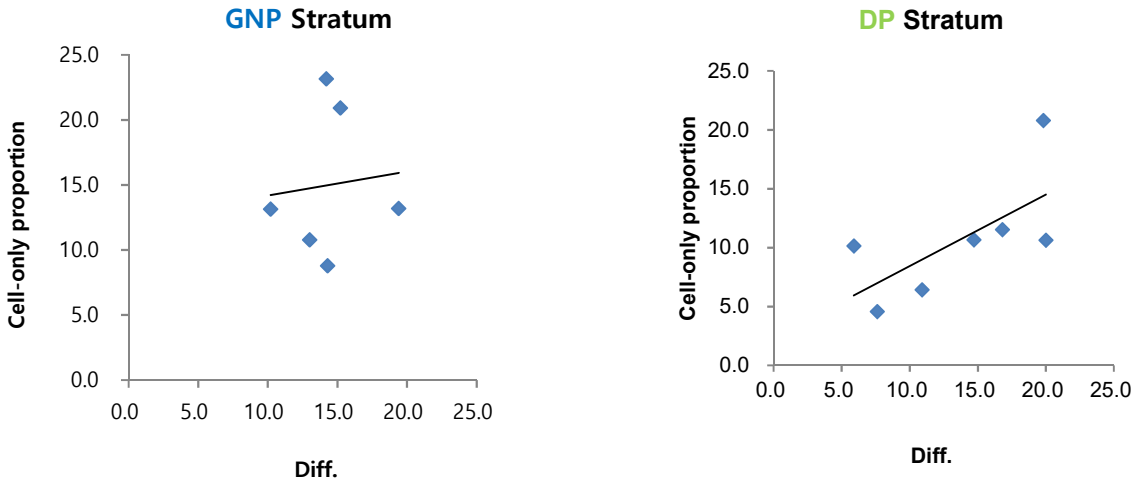
Stratum	Listed Landline	Landline RDD	Cell RDD	Dual RDD	Actual Result
GNP	54.6%±8.7	53.9%±6.0	48.5%±6.0	49.9%±4.8	
Diff. (p.p.)	**-0.3**	**-1.0**	**-6.4**	**-5.0**	54.9%
N	191	385	254	639	
DP	65.7%±11.2	61.7%±8.8	61.8%±9.6	57.6%±7.7	
Diff. (p.p.)	**9.3**	**5.3**	**5.4**	**1.2**	56.4%
N	98	162	106	268	
Others	59.9%±19.1	49.9%±13.7	50.8%±15.4	50.3%±11.6	
Diff. (p.p.)	**9.5**	**-0.5**	**0.4**	**-0.1**	50.4%
N	46	77	40	117	

Note. See the note in Table 1.

Data presented in Table 2 show the distributions of expressed interest in politics and discussion of the campaign with others among self-reported voters in the three subsamples and combined in the dual-frame sample. Of course, there is no external source like the census or the KNEC study to which equivalent measures can be compared to provide a reference point;

we can only compare the responses observed in the three different subsamples. For these measures, respondents from the cell phone RDD sample were more likely to express an interest in politics (be "Very" or "Somewhat" interested) and to say that they talked to others about the election before voting than those in either of the landline subsamples. When the three subsamples were combined through weighting, levels of interest and discussion were higher than among the respondents in the listed landline subsample alone, equivalent to the typical sampling frame for the pre-election polls. These results were also replicated when looking at those who reported they voted for the winner.

Next, we take a look at the results by the three geopolitical strata used in the sample design. Data presented in Table 3 show the proportions who said they voted for the winner in their area designated as the "GNP stratum," "DP stratum," and "Others stratum." When compared to the actual results in the final column, in the GNP stratum, the respondents consistently underreported in all frames, especially for the cell RDD and dual RDD. In the DP stratum, there is some evidence of social desirability, since they consistently overreported in all frames. It is especially noteworthy that the difference (Diff.) between the estimate of the vote share and the actual result in the GNP stratum is very small in the listed landline and landline RDD samples, at just -0.3 and -1.0 percentage points, respectively.



(a) r = 0.10

(b) r = 0.67

Figure3. Correlation between the Difference in Pre-election Poll Estimates and Proportion of Cell Only Population by Units within Each Party Stratum

Why should this happen even though the listed landline and landline RDD samples should have a noncoverage problem? The answer is given in Figure 3 where we show the scatterplots of correlations between the differences from actual results of pre-election poll estimates and the rates of cell phone only ownership for the areas in each of the two strata where there were major party winners.[2] In the DP stratum, the correlation between the difference in polls and the proportion (percent) of the cell only populations not covered by landline frame is very high (r = .67), while it is very low (r = .10) in the GNP stratum. This suggests that the estimation errors in the GNP stratum were not very dependent on the undercoverage in the landline frame due

to the cell only population, but this undercoverage was much more important in the DP stratum where in fact the underestimates of support in the pre-election polls were substantial.

As a way to validate this relationship, we turn to the two areas in Korea – Seoul and Incheon – where we have moderate subsample sizes sufficient to estimate support for the winning party. Because the post-election survey asked voters whether they supported the winner in their constituency, we can take advantage of the fact that the GNP won by a very narrow margin in Seoul while the DP won by a comfortable margin in Incheon even as the pre-election polls showed the GNP ahead by substantial margins in both.

Table4.
Self-reported Vote for the Winner by Frame for Reported Poll Results in the Week Before the Election in Seoul and Incheon Compared to the Election Results.

Seoul

		Actual Result	5 Media Groups					
			Poll A	Poll B	Poll C	Poll D	Poll E	Average
Pre-Election Poll	GNP	47.4%	48.9%	46.7%	50.8%	50.4%	51.6%	2.3
	Diff.		1.5	-0.7	3.4	3.0	4.2	
	N		507	1000	800	1000	806	
			Listed Landline		Landline RDD	Cell RDD	Dual RDD	
Post-Election Survey	GNP	47.4%	55.1%±16.8		50.6%±10.8	41.8%±10.6	45.2%± 8.3	
	Diff.		7.7		3.2	-5.6	-2.2	
	N		50		116	85	201	

Incheon

		Actual Result	5 Media Groups					
			Poll A	Poll B	Poll C	Poll D	Poll E	Average
Pre-Election Poll	DP	52.7%	33.4%	34.4%	34.6%	32.9%	31.8%	-19.3
	Diff.		-19.3	-18.3	-18.1	-19.8	-20.9	
	N		506	1000	500	800	803	
			Listed Landline		Landline RDD	Cell RDD	Dual RDD	
Post-Election Survey	DP	52.7%	55.2%±33.2		43.9%±23.3	53.8%±20.4	52.7%± 16.8	
	Diff.		2.5		-8.8	1.1	0.0	
	N		13		26	25	51	

Note. See the note in Table 1.

Data presented in Table 4 show this to be the case. The pre-election polls suggested the GNP had a substantial lead over the DP in Seoul, but the actual outcome was very close as they won by only .6 percentage points. The pre-election polls on average overestimated their support by

2.3 percentage points, and individually by no more than 4.2 percentage points. Overall, the dual-frame sample in the post-election survey underestimated the GNP support as the winner by 2.2 percentage points, within the margin of error. Since the pre-election estimates of GNP support were generally accurate, the dual-frame sample does not provide much improvement. The picture is different in Incheon, where the DP won the election despite the pre-election polls seriously underestimating its support by an average of 19.3 percentage points. Because the dual-frame design improved greatly estimates of the DP support due to its inclusion of cell phone only respondents, the estimate of their support was right on as a function of a slight overestimation in the listed landline subsample, an underestimate in the landline RDD subsample, and an accurate estimate in the cell phone RDD subsample.

Conclusions

This analysis suggests that the 2010 pre-election polls in Korea misestimated the actual results due to errors of nonobservation such as noncoverage, sampling errors, and nonresponse. The pre-election polls did not indicate a spiral of silence effect as the support for the different parties did not vary over time across the different polls. Furthermore, the pattern of differences from the actual outcome did not suggest that the second-place finisher was always underestimated. Our findings suggest that the problems of non-random respondent selection and the omission of cell phone only respondents seem to be the primary sources of error. Based on our results, we believe that adding a cell phone RDD sample would be beneficial to estimating election results in pre-election polls. However, the use of a cell phone sample would not completely eliminate response bias. Random selection of respondents within households in place of the use of quotas would also improve estimation.

The analysis presented here focused primarily on methodological explanations for estimation errors in the 2010 pre-election polls in Korea. But they have general applicability to the current issues of appropriate sample frames and contact strategies that all pre-election pollsters face. In the 2012 election in the United States, for example, the overall performance of the pre-election polls was generally good. However, the Gallup Poll estimated that Mitt Romney would prevail over Barack Obama, and he did not. In their post-election evaluation of their methodology (Gallup 2013), Gallup acknowledged that their use of a listed landline frame instead of an RDD list-assisted landline frame created problems for them, and they have reverted to the RDD frame.

The consequence of these errors in Korea was likely to mislead citizens about the campaigns' status and the candidates' standing in most of the constituencies, especially since news organizations sponsored the majority of the polls and widely distributed the results through their campaign coverage. It is also possible that a consequence of these estimation errors may have been to create expectations in the public that in turn led to bias in the answers respondents gave in subsequent polls.

The methods used by Korean pollsters have changed since the 2010 election, but they still do not meet the accepted standards of RDD sampling by combining landline and cell phone frames and the random selection of individual respondents for their surveys. Our results suggest that estimation would be improved substantially by incorporating both of these changes into pre-election polling, as well as improving weighting procedures for the data.

Additional research is needed to understand human factors that may contribute to a significant social desirability component in party preferences that respondents offered. This suggests the need for a panel component in some pre-election polls so it is possible to see

whether initial expressions of preference for candidates from second or third-place parties become preferences for the leading party in subsequent interviews, as well as knowledge of who was leading and who eventually won. With such a sophisticated design it will be possible to distinguish the relative contributions of survey methods and social phenomenon to the explanation of misestimation errors in pre-election polls.

FOOTNOTES

[1] Response rates were computed using the formulas suggested by the American Association for Public Opinion Research (AAPOR 2011). The RR1 (RR5) rate was 34.1% (69.9%) for the landline sample, 27.4% (84.7%) percent for the cell sample, and 31.0% (75.2%) percent for the combined dual-frame sample.

[2] The pre-election polls were jointly conducted by the three major broadcast networks during May 24-26 (e.g., see Figure 1 & Figure 2). The larger difference of differences between the actual results and the reported poll results for the candidates in each area was used to scatterplot. The estimates of the cell phone only population in each area were obtained from an analysis of the raw data from the Survey on Internet Usage (2009) conducted by the Korea Internet and Security Agency.

ACKNOWLEDGEMENTS

We would like to thank the followings for assistance: Sung-Joon Hong, Hee-Jin Kim, Young-Je Woo, Eun-Hee Choi. The Nam Center for Korean Studies at the University of Michigan provided travel support for collaboration in preparing this manuscript, for which Traugott is grateful. An earlier version of this paper was presented at the 64th Annual Conference of the World Association for Public Opinion Research in Amsterdam in 2011.

REFERENCES

Atkeson, L. R., Adams, A. N., Bryant, L. A., Zilberman, L., & Saunders, K. L. (2011). Considering mixed mode surveys for questions in political behavior: Using the internet and mail to get quality data at reasonable costs. *Political Behavior, 33*, 161-178.

Bolstein, R. (1991). Predicting the likelihood to vote in pre-election polls. *Journal of the Royal Statistical Society, 40*, 277-283.

Choi, E. H., Kim, S. W., &Cho, C. H. (2011). Coverage and demographic characteristics of telephone households in Korea, Paper presented at *the Spring Conference of the Korean Association for Survey Research*, Seoul.

Choi, E. H., Kim, S. W., Cho, C. H., & Couper, M. P. (2011).Comparing cell phone and Internet coverage in the USA and South Korea. Paper presented at the Joint Statistical Meetings, Miami Beach, FL.

ChosunIlbo (2010a).*http://www.chosun.com/site/data/html_dir/2010/05/17/2010051702917.html*

Cochran, W. G. (1977). *Sampling techniques*. New York: John Wiley and Sons, Inc.

Crespi, I. (1988). *Pre-election polling: Sources of accuracy and error*. New York: Russell Sage.

Dong-AIlbo (2010a).*http://news.donga.com/view.php?id=Print_Donga|3|20100518|28422824|1*

Durand, C. (2011). Polls at the subnational level: The Canadian case. Paper presented at *the 64th annual conference of the World Association for Public Opinion Research,* Amsterdam.

Durand, C., Foucault, M., Goyder, J., & M. Deslauriers. (2010). Mispredictions of electoral polls: A metaanalysis of methodological and sociopolitical determinants over 50 Years. Paper presented at *the International Sociological Association World Congress*, Gothenburg, Sweden.

Freedman, P. & Goldstein, K. (1996).Building a probable electorate from preelection polls: A Two-Stage Approach. *Public Opinion Quarterly*, 60, 574-587.

Gabler, S., & Häder, S. (2002). Idiosyncrasies in telephone sampling – the case of Germany. *International Journal of Public Opinion Research, 14(3),* 339–345.

Gallup Organization. (2013). "Gallup 2012 Election Polling review." Available at http://www.gallup.com/poll/162887/gallup-2012-presidential-election-polling-review.aspx (accessed September 20, 2013).

Groves, R. M. (1989). *Survey errors and survey costs*. New Jersey: John Wiley & Sons, Inc.

Hopkins, D. J. (2009). No more wilder effect, never a Whitman effect: When and why polls mislead about black and female candidates. *Journal of Politics,71,* 769-781.

JoongangIlbo (2010). *http://koreajoongangdaily.joinsmsn.com/news/article/option/rticle_print.aspx*

Kalsbeek, W. D., & Agans, R. P. (2008).Sampling and weighting in household telephone surveys. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. D. Leeuw, L.Japec, & P. J. Lavrakas, *et al*. (Eds.), *Advances in telephone survey methodology* (pp. 29-55). New Jersey: John Wiley & Sons, Inc.

Karp, J. and Brockington D. (2005).Social desirability and response validity: A comparative analysis of overreporting voter turnout in five countries. *Journal of Politics 67,* 825-840.

Keeter, S. (2006).The impact of cell phone noncoverage bias on polling in the 2004 presidential Election. *Public Opinion Quarterly, 70,* 88–98.

Keeter, S., C. Kennedy, A. Clark, T. Thompson, & Mokrzycki, M.(2007). What's missing from national landline RDD surveys? *Public Opinion Quarterly, 71,* 772-792.

Kim, S. W., Lee, S. K, Hong, S. J., & Park, S. H. (2012).List-assisted RDD sampling in Korea: Testing the feasibility of national survey under within-household selection. *International Journal of Public Opinion Research, 24(1),* 79-92.

Kim, S. W., Park, S. H., & Hong, S. J. (2011).The development of alternative dual RDD frame telephone survey designs in Korea: theory, practice and challenges, Paper presented at *the Spring Conference of the Korean Statistical Association*, Daejeon.

Korean National Election Commission. (2010). *http://nec.go.kr/nec_new2009/InsStatisticData.do?module=NEC*

Lohr, S. L. (1999). *Sampling: Design and analysis.* North Scituate, MA: Duxbury Press.

Moreno A., Aguilar, R., &Romero, V. (2011). The unnatural left-right coalitions: challenges for preelection polls in Mexico. Paper presented at *the 64th annual conference of the World Association for Public Opinion Research*, Amsterdam.

Nicolaas, G., & Lynn, P. (2002).Random-digit dialing in the UK: Viability revisited. *Journal of the Royal Statistical Society, Series A, 165, Part 2*, 297–316.

Noelle-Neumann, E. (1974). The spiral of silence: A theory of public opinion. *Journal of Communication, 24*, 43-51.

Park S. H., Kim, S. W., Traugott, M., & Lee, S. K. (2011). An enhanced approach for solving the overlap problem in dual frame RDD surveys. Paper presented at *the Joint Statistical Meetings*, Miami, Florida.

Sala, E., &L. Fumagalli. (2010). Coverage error and Italian polls: The case of the 2006 general elections. Paper presented at *the World Congress of Sociology*, Gothenburg, Sweden.

Smith, T. W. (1978). In search of house effects: A comparison of responses to various questions by different survey organizations. *Public Opinion Quarterly, 42(4),* 443-463.

Smith, T. W. (1982). House effects and the reproducibility of survey measurements: A comparison of the 1980 GSS and the 1980 American national election study. *Public Opinion Quarterly, 46(1),* 54-68.

Statistics Korea (2010).*http://kosis.kr/abroad/abroad_01List.jsp?parentId=A Census*

Traugott, M. W., R. M. Groves, & J. M. Lepkowski. (1987). Using dual frame designs to reduce nonresponse in telephone surveys. *Public Opinion Quarterly, 51*, 522-539.

Traugott, M. W. and C.Wlezien. (2009). The dynamics of poll performance during the 2008 presidential nomination contest. *Public Opinion Quarterly, 73, No. 5 2009,* 866–894.

BIOGRAPHICAL NOTES

Sun-Woong Kim is the director of the Survey & Health Policy Research Center and a professor at the Department of Statistics at Dongguk University, Seoul, South Korea. His research focuses on the dual frame RDD landline and RDD cell sample designs. He has also worked on model-based sampling and public-use software for sample selection.

Michael W. Traugott is a professor of communication studies and political science and senior research scientist at the Center for Political Studies in the Institute for Social Research, University of Michigan.

So-Hyung Park is a PhD at the Department of Statistics, Dongguk University, and has served as a senior researcher for the same university's Survey & Health Policy Research Center.

Sang-Kyung Lee is the chief executive officer at the Hyundae Research Institute, Seoul, South Korea.