

# **Poll Accuracy Measures in a Quasi-Two-Way Election: An Empirical Assessment of Taiwan's 2000 Presidential Election**

Chien-chou Su

Associate Professor

Department of Communications Management, Shih Hsin University

Mandy Sha

Research Survey Methodologist

RTI International

## **ABSTRACT**

There is little systematic study on the question of how poll accuracy should be measured. Should different measures be adopted in different election situations? Expanding upon Mitofsky and Traugott's assessment of the 1996, 2000 and 2004 U.S. pre-election polls, this paper assesses democratic Taiwan's polls in its second direct presidential election in 2000. In this paper, we first discuss the feasibility of accuracy measures in different election situation, and then focus on designing a suitable poll accuracy measurement for Taiwan's 2000, a quasi-two-way election. Next, we adopt the proposed accuracy measurement as a response to evaluate the Taiwanese poll performance. Finally, we analyze the multitude of factors that influence polling accuracy by using the proposed and modified measures to compare the effects of eight factors.

**Key Words:** poll accuracy measure, poll performance, quasi-two-way election

## I. Introduction

Discrepancies between similar surveys undermine the credibility of the results, especially when they share coverage of a population, and are conducted around the same time. Due to a lack of actual feedback, they often cannot be evaluated. However, pre-election polls offer a unique advantage that they can be examined by the final outcome of the election. Measuring polling accuracy will contribute greatly to the public's knowledge about their reliability, and serve to encourage self-regulation in the polling industry. Through measuring the margin of error from the final vote tally, common retrospective assessments post-election can be made. First, pre-election polls can be reexamined from a particular election (Traugott, 2001: 389-419; Mitofsky, 1998: 230-249; NCPP, 2007a) or from a historical comparison (Mitofsky, 1998: 230-249; Garand and Parent, 1991: 1011-1031; Buchanan, 1986: 222-227; NCPP, 2007a). Second, influential factors for poll accuracy can be investigated (Lau, 1994: 2-20; Converse and Traugott, 1986: 1094-1098; Crespi, 1988: 25-33). A variety of poll error calculation methods have been used for measuring accuracy in the past and each serves a different purpose. Here rises a significant problem: What are the criteria used to determine a suitable polling accuracy method?

Expanding upon Mitofsky (1998: 230-249), Traugott's (2001: 389-419), and Traugott's (2005: 642-656) assessment of the 1996, 2000 and 2004 U.S. pre-election polls, this paper examines democratic Taiwan's polls in its second direct presidential election in 2000. Data is obtained from 68 polls conducted by 22 sponsoring/polling organizations and published in Taiwan's major newspapers and magazines during the two months preceding the election.<sup>1</sup> Poll data used exclude partial results from tracking polls or those polls that only covered a sub-demographic group. First, we discuss the feasibility of accuracy measures in different election situation, and then focus on designing a suitable accuracy measurement for Taiwan's 2000 polls. Next, we adopt the proposed accuracy measurement as a response to evaluate the Taiwanese poll performance. Since circumstances surrounding major elections can be dynamic as evident in Taiwan's 2000 election, we restrict our

---

1 January 8, 2000 to March 7, 2000. Among the 22 organizations, some released more than one polls, thus totaling 68.

assessment to 22 final polls released close to Election Day.<sup>2</sup> Finally, we analyze the eight factors that influence polling accuracy by using the proposed and modified measures: sample size, percent undecided, number of days leading to the election, types of sponsoring organizations, sponsorship, frequency of polls published, days in the field for polling data collection, and weekdays only field period. We used results from all 68 polls in the initial analysis, and then removed 25 outliers in the ensuing in-depth analysis. A historical poll performance comparison can not be made because 2000 was only Taiwan's second presidential election.

## II. Presidential Election in 2000 and Pre-Election Polling in Taiwan<sup>3</sup>

The Taiwanese polling experience in 2000 presidential election distinguishes itself from that of the United States in several ways. First, about 36% of polls erroneously foretold the victory of a candidate who actually came in third in the election and lagged behind both leading candidates by more than 13 percentage points. Second, the majority of polls showed an “undecided” percentage of more than 25 percent as reported by the volatile

---

2 Among the 22 final polls, only 5 were conducted more than 20 days away, as shown in Table 2 in the body of the paper. We are not able to use final polls conducted the day before the election because Taiwanese election law prohibits polling information to be published starting the 10<sup>th</sup> day before the election. Half of the polls we assessed were released either one or two days before the ban, March 8.

3 Following the communist victory in the Chinese civil war, the defeated KMT (KuoMinTang, or Nationalists) fled Chinese mainland to Taiwan in 1949. China has since declared Taiwan as a renegade province to be reunited into sovereign Chinese soil. It has also effectively barred Taiwan from joining any major international organization of consequence (with the exception of its recent membership in WTO). Furthermore, Taiwan is not diplomatically recognized by any major industrialized country, including the United States, nor by the United Nation. However, the island has prospered from an agrarian economy of rice and sugar to become one of East Asia's economic “dragons,” and has suffered little compared to its counterparts in the Asian financial crisis in 1998–1999. According to the World Fact Book published by US's Central Intelligence Agency, Taiwan's foreign reserves are the world's third largest (CIA, 2007). Democratization movement in Taiwan formally commenced in the late 80s when the opposition force established the DPP (Democratic Progressive Party) that holds a clear stance on promoting Taiwan's independence from China. Shortly after, the ruling KMT ended the 40-year-long martial law.

electorate. Furthermore, Taiwan saw a recent phenomenon of a proliferation of polls cited in the media and used by the campaigns. On the contrary, US pollsters have been releasing polls for more than 50 years, the small percentage of undecided voters is of little interest to the public, and no polls in the United States predicted a winner that ran third in the election outcome. Although these interesting contrasts exist between the new and veteran polling experiences, both Taiwan and the US face the same dilemma of predicting the election result and measuring poll performance with fairness and precision.

Pre-election polls were of great interest to the consumers of predictions in Taiwan's 2000 presidential election. More than 300 polls<sup>4</sup> were published throughout the campaign and received significant media coverage. However, they were conducted by both credible and problematic sources: research institutes, news organizations, or conveniently established polling agencies. Some organizations sponsored a survey house (polling organization) to conduct polls that the sponsoring organization later released,<sup>5</sup> but in most cases, an organization did not sponsor or partner with another survey house. Instead, they conducted the polls themselves. As part of the campaign, the three leading presidential candidates also released poll results daily. While they accused each other of using "fake" polls in an attempt to mislead voters, all claimed to lead the polls. The election law restricts poll results to be published 10 days before the election. Consequently, an unusually high number of polls were released the day before the ban came into effect.

To some observers, Taiwanese politics has divided along ethnic lines. The former ruling party KMT has been run by Chinese mainlanders that

---

4 For more than 20 years, pollsters in Taiwan have been publishing predictions of metropolitan election results in academic journals such as the *Chung Guo Statistics Journal*. As the ban on private newspaper publishing was lifted in 1987, polls have proliferated in the mass media and the internet (Huang, 2004). International pollsters reported in mid-90s that polls in Taiwan had generally concerned social and economic issues rather than political questions. However, pre-election polling has emerged rapidly in the late 90s. Public opinion polls have been adopted by the two major political parties, KMT and DPP, as the basis of their candidate nomination system in the party primaries (Wu and Fell, 2001: 23-45).

5 For example, Global TV and the Journalist sponsored CNA to conduct polls (Global TV/the Journalist/CAN), and SET TV sponsored to conduct polls (SET TV/Gallup) in Taiwan's 2000.

migrated to Taiwan in 1949, while 80 percent of Taiwanese population is ethnic Taiwanese. It made history in 1996 when Lee Teng-hui of KMT became the first native-born president in Taiwan's first democratic direct election. When the second direct presidential election was held in 2000, it quickly developed into a charged race led by three main candidates: Lien Chan of KMT, Chen Shui-Bian of DDP, and an independent candidate James Soong who departed from the KMT to launch his own campaign.

The stakes of Taiwan's 2000 were so high that political strategists and the mass media used a term, "dump-save effect," which is also known as "strategic voting" or "sophisticated voting" (Alvarez and Nagler, 2000: 57-75; Schoen, 1999: 473-496; Blais and Nadeau, 1996: 39-52; Bowler and Lanoue, 1992: 485-499; Burden, 2005: 603-618), to describe a psychological factor that drove electoral behavior. When the voters saw that their preferred candidate has little chance of winning, they might choose to "dump" him and "save" (vote) for the second favored candidate against the threat of the less-preferred candidate. "In cases where there are three parties operating under the simple majority, single ballot system, the electors soon realize their votes are wasted if they continue to give them to the third party; whence their natural tendency to transfer their vote to the less evil of its two adversaries in order to prevent the success of the greater evil" (Duverger, 1954: 55-56).

While research shows that there are many factors influencing voter behaviors, the campaigns in this election exposed Taiwanese voters to different "dump-save" scenarios. For instance, KMT urged its ethnic Chinese mainlander constituency to "dump Soong, save Lien" in fear that the Taiwan-born Chen would win. On the other hand, DPP promoted the sense of "dump Lien, save Chen" against the less-preferred mainland-born Soong who does not support Taiwanese independence from China. In addition, voters who had grown tired of the KMT monopoly and would not want Lien to continue the legacy might dump either Chen or Soong to save the stronger candidate. However, the volatile Taiwanese electorate expressed itself in the pre-election polls with a high "undecided" percentage, usually more than 25%.

Soong enjoyed a solid lead in all polls until four months before the election. Then the Chung Hsing Bills Finance Corporation scandal broke out. The KMT alleged that Soong embezzled party funds of nearly US \$32 million while he was a high-ranking party leader. At that time, most polls indicated that it was a turning point for Soong. TVBS polls showed that

Soong's support dropped more than 10%. The voters had apparently been absorbed by Chen because Lien's numbers did not climb. Soong remained a formidable challenger to Chen. Soong and Chen remained neck and neck in the polls.<sup>6</sup> Although Lien consistently trailed in the local polls, almost all of the election polls over-estimated his votes in the election. Significant news events continued as the Election Day approached, including the Chinese verbal military threat, multiple big political rallies, and the endorsement of Chen by the respected Taiwanese Nobel laureate, Yuan-Tseh Lee, who headed Academic Sinica. Together the three candidates garnered more than 99% of the final voter turnout: Chen won the presidency with 39.3% of the popular vote and marked an end to 54-year KMT leadership, and the independent candidate James Soong lost by a close margin and fared second with 36.8%. Although the campaign saw three major candidates, the results showed that this was a two-way election after all. Lien of the KMT placed a distant third at 23.1%. Usually, the sample size of pre-election polls are ranged from 800 to 1200, that means the sampling errors for candidates' support difference are 7% around for 95% confidence. In the study, such an election with three major candidates, where the popular votes difference for leading two are too close (less than 7%) to call for regular polls, while the third has a significant difference (more than 10%) behind them is called as a "quasi-two-way election".

### III. Polling Accuracy Measurement

#### I. Measures

The Social Science Research Council (SSRC) considered eight possible methods for measuring poll error following the fiascos of the 1948 polling experience in the US. They noted that each of these methods has been used at various times and for various purposes, and each has advantages and disadvantages (Mosteller et al., 1949: 54-59). The SSRC committee's definitions are as follows.

Measure 1: The difference (without regard to sign) in percentage points between the percentage the winning candidate received in the election and what he received in the poll.

Measure 2: The difference in percentage points between the leading

---

6 The scandal broke out four months before the election, and the polls assessed in this paper were conducted afterwards that reflect the tight race between Soong and Chen.

candidates' share of the major party vote from a poll and from the actual vote.

Measure 3: The average (without regard to sign) of the percentage point deviation for each candidate between his estimate and the actual vote.

Measure 4: The average difference (without regard to sign) between a ratio for each candidate and the number one, where the ratio is defined as a candidate's estimate from a poll divided by the candidate's actual vote.

Measure 5: The difference between two differences, where the first difference is the estimate of the vote for the two leading candidates from a poll and the second difference is the election result for the same two candidates.

Measure 6: The maximum difference in percentage points between a party and the actual vote.

Measure 7: The chi-square to test the congruence of the estimated and actual vote distributions.

Measure 8: The difference between the predicted and actual electoral vote.

Mitofsky (1998: 230-249) observed that Measures 1, 2, 3, and 5 are commonly adopted by the polling community, while Measures 4, 6, 7, and 8 were not considered as viable options. Further, he argued that the best choice of polling accuracy measure appears to be between Measures 3 and 5. After examining the two measures, Mitofsky advocated using Measure 5 because it allows a suitable comparison over time for both two-way and three-way elections. He also pointed out two other advantages. First, it eliminates the need to allocate "undecided" for polls that do not do so. Secondly, Measure 5 has become the measure that the public is accustomed to hearing about as point spread or margin, are often used for reporting poll results. Nonetheless, Mitofsky used Measures 3 and 5 throughout his evaluation of the 1996 US presidential election polls. Traugott (2001: 389-419) and Traugott (2005: 642-656) subsequently followed upon Mitofsky's analysis and used the same two measures in their assessment of the 2000 and 2004 US presidential election, respectively.

Other scholars and researchers have preferred and recommended different measures. In 2000, the National Council on Public Polls (NCPP) presented calculations of presidential poll error from 1936 to 2000 based on Measure 5. Crespi (1988: 102-110) concluded that Measures 1, 3, and 6 should be used because empirical data from 343 polls showed that these three measures produce very similar results with highly positively correlation. Furthermore, Crewe (1997: 569-585) used Measure 3 for his analysis of 1992 British

election and called it “the true test of poll’s accuracy.” Martin, Traugott, and Kennedy (2005: 342–369) have proposed a new measure of predictive accuracy based upon the natural logarithm of the odds ratio of the outcome in a poll and the outcome of the election. The new measure is applied to summarize the results of three presidential elections (1948, 1996, and 2000) polls.

In fact, the estimated votes of candidates are multinomially distributed. The measurement of polling accuracy extracts a “single-dimensional index”, which is polling error from the congruence level of the estimated. Therefore, the flexibility of the measure should take into account the final vote distribution of the candidates in the election. Additionally, a good polling accuracy measure has to connect with the public interest. Naturally, different measures should be used in different election situations.

In the following, we classify the SSRC methods according to the election status, which is categorized by the number of major candidates (one-way, two-way, multi-way) in the election result. Here the major candidates are indicated as those who have a significant lead over other candidates in the election.

1. One-Way (Land-Slide) Election: Measure 1: As indicated by Mitofsky (1998: 230–249), “while simple and easily understood, [SSRC Measure 1] is artificial unless the test in the election is whether the leading candidate gets 50 percent or more. The number of the leading candidate alone is of little value in describing the status of an election.” For example, suppose that an election situation that was 40 percent for A, 30 percent for B, and 30 percent for C, while a poll showed 40 percent for A, 50 percent for B, and 10 percent for C. Measure 1 would have a zero error, although the poll predicted a wrong winner. In other words, SSRC Measure 1 is meaningful only when a poll predicted the correct winner. Measure 1 is particularly useful in a landslide election. Here, we define an election as a landslide if the winner got more than 60% of the vote. Obviously, polls need to pick the correct winner in order to be judged as accurate. Instead of measuring the margin, or average error, between the leading candidates, the percentages of votes the winner got seems to meet the public’s interest more closely. Therefore, Measure 1 might be a good choice for assessing poll’s accuracy in such an election situation.

2. Two-Way Election: Measures 2 and 5: Measure 2 recalculates in percentage the two leading candidates and adds to 100 percent. In fact, Measure 2 is identical to Measure 5 with the undecided and others proportionate-



ly allocated. Nevertheless, Measure 5 (also Measure 2) alone is of little value if we have no information about whether or not the two leading candidates are correctly predicted. For example, suppose that election situation was 40 percent for A, 37 percent for B, and 23 percent for C, while a poll showed 33 percent for A, 30 percent for B, and 37 percent for C. Measure 5 would have a zero error, although the poll incorrectly predicted the two leading candidates, and the winner was predicted as the third. Therefore, when Measures 2 or 5 are adopted, we need an auxiliary mechanism to prevent the wrong measurement as that committed by the poll illustrated above. Measures 2 and 5 are particularly useful in a two-way election. However, polls need to pick correctly the two leading candidates in order to be judged as accurate in such an election.

3. Multi-Way Election: Measures 3, 4, 6, and 7: The advantage of the polling accuracy measures in this category is that they use the most complete information. However, Measure 6 concerns the largest error for any candidate, even a minor candidate who weighs little in the election outcome. It is therefore not considered a viable option. Measure 4 is an average of relative error, while Measure 3 is an average of pure error. Both measures tend to respectively underestimate or exaggerate small percentage point differences induced from minor candidates. For example, suppose an election situation that was 50 percent for A, 49 percent for B, 0.5 percent for C, and 0.5 percent for D, while a poll showed 45 percent for A, 54 percent for B, 1 percent for C, and 0 percent for D. Measures 3 and 4 would have an error of 2.75 and 55 percentage points, respectively. However, while the leading candidates A and B took 99 % of the vote, their average error sums up to be 5% and 10% as respectively calculated by Measures 3 and 4. The numbers for minor candidates greatly influenced the error calculation. Measures 3 and 4 are only particularly useful if the candidates share approximately equal votes. Crespi (1988: 6-7) suggested including candidates in the calculation only if they received more than 15 % of the vote. Similarly, to compare with the two-way election situation, we suggest including those candidates who do not have a clear-cut (more than 15%) difference with the leading candidates in calculating Measures 3 and 4.

Rather than a poll error measure, Measure 7 is a Chi-square statistic for testing the congruence of the multinomial distribution. It can be used as an index of testing the significance of the fact that the estimated votes and actual votes are identical multinomial distributed. The larger the Chi-square value is, the worse the performance of the poll is. Here we suggest

include all candidates that received more than 0.5 % of the vote.

According to the above classification, US presidential elections from 1936 to 2000 suggest the use of a two-way election measure, except in four elections where the winner had a significant lead over the second place. They occurred in 1984 when Reagan gained 59% votes (18% lead ahead of Mondale), 1972 when Nixon gained 61% votes (23% lead ahead of McGovern), 1964 when Johnson gained 61% votes (22% lead ahead of McGovern), and 1936 when Roosevelt gained 61% votes (24% lead ahead of Landon). In these elections, the leading two candidates gained more than 98 % votes that made them either a true two-way election or one that is similar. This means we actually do not need to address the issue of polling accuracy measures. We endorse Mitofsky's (1998: 230-249) conclusion: Measure 5 should be the best choice for comparing US presidential elections over time.<sup>7</sup>

## 2. Minor Parties and Undecided Votes

Besides determining a suitable polling accuracy measure, handling the estimates of minor parties and the undecided vote poses a significant problem. Since there is no undecided category in the election result, it is necessary to allocate them to insure the comparability from polls to polls. Crespi (1988: 42-43) suggested a "15 percent rule" that a minor party candidate who receives less than 15 percent of vote should be eliminated from the poll error calculation. Mitofsky (1998: 230-249) discussed the four allocation methods that address the undecided category:

1. Allocate the undecided in proportion to the votes for candidates in a poll.
2. Allocate the undecided evenly between the two major parties.
3. Allocate the undecided to the challenger, if there is an incumbent.
4. No allocation is required or needed.

Crespi (1988: 6-7) advocated the use of method 1, proportional allocation, which in effect assumes that the undecided do not vote on the Election Day or split proportionately as the decided in a poll. Mitofsky (1998: 230-249) and Traugott (2001: 389-419) also suggested the use of proportional allocation before calculating Measure 3. On the other hand, they indicated that no undecided allocation is needed in calculating Measure 5. That means

---

7 It should be noted here that in a two-way election, the distribution of the estimated vote is reduced to a binomial distribution. The effect is that Measures 1, 2, 3, 5, and 6 then become identical.

allocation method 4 is used for Measure 5. In fact, the allocation scheme can be further interpreted as follows: If there is no allocation used in calculating Measure 3, it is identically mapped to when the undecided are all allocated to those non-major candidates. We can easily see that it is very unreasonable. Similarly, if there is no allocation used in calculating Measure 5, it is identically mapped to a scheme allocating the undecided evenly, not proportionally to the two leading candidates. We can easily understand that this scheme is not fair, except when the vote difference of the two leading candidates is tiny. Therefore, we will proportionately allocate the undecided and minor party votes to the leading two candidates before calculating Measure 5.

## **IV. Assessing Taiwan's 2000 Pre-election Polls**

### **I. Establishing a Suitable Measure**

Lien, who came in third in the election, gained almost a quarter of total votes, but lost by a clear-cut difference with the two frontrunners, Chen and Soong. The three of them together make up more than 99% of the vote. The campaign saw three main leading candidates, but according to the classification rule, it shows that this was a two-way election after all as Lien of the KMT placed a distant third, more than 13% behind Chen and Soong. Taiwan's 2000 is a two-way election with a potential third, it is called as a "quasi-two-way election" throughout the paper. Measure 5 should be superior to Measure 3 for Taiwan's 2000, a quasi-two-way election. However, some modifications are necessary for the following reasons. First, more than one-third of Taiwanese final polls predicted a winner that actually came in third, while US polls committed no such mistakes. Additionally, generally more than 25 percent of undecided respondents appeared in Taiwanese pre-election poll reports. In the US, however, the undecided were usually allocated by the pollsters, or less than 5 percent when the percentage were actually shown (NCPP, 2007b).

Modified Measure 5: Mitofsky (1998: 230-249) advocated the use of SSRC Measure 5 because it eliminates the need to allocating undecided and minor party supports. He argued that under the premise that the goal of pre-election polls is forecasting, the responsibility for allocation should rest with the pollster because the public does not have the information nor the technical knowledge necessary for sophisticated allocation. However, the goal of pre-election polls in Taiwan seemingly aimed not to predict the election outcome, but to inform the public of the likely winners. The polls

always revealed a high percentage of undecided, which was not allocated by pollsters. In fact, only two of the sixty-eight Taiwanese polls examined in this paper did their own allocation.

Measure 5 calculates the difference between two differences, where the first difference is the estimate of the vote for the two leading candidates from a poll and the second difference is the election result for the same two candidates. Table 1 compares the second difference and the first difference for two real polls conducted jointly by two news organizations, Global and the Journalist on January 24th, and a polling house called Poll Survey Organization on March 3rd. To further clarify the comparison, a pseudo (dummy) poll “Extreme Case” displays an unreasonable scenario. As shown in Table 1, the second difference is 2.46% (from a total 76.14%), while the first difference is 0.3% (from a total 46.3%), 2.5% (from a total 45.1%), and 2.46% (from a total 22.46%) for Global/the Journalist, Poll Survey Organization, and the Extreme Case.

Without allocating the “undecided”, the third candidate Lien, and other minor candidates’ supports, the error measure calculated for Measure 5 is highest for Global/the Journalist 2.16, but dropped significantly to 0.04 for Poll Survey Organization, and 0 for the Extreme case. However, it is meaningless to subtract the second difference from these first differences, since their total percentages are quite different. Exaggeratedly, the error calculation of SSRC Measure 5 is zero but meaningless.

**Table 1 Comparison of Measure 5 and Modified Measure 5**

Organizations and Case	Coverage Date	Soong (Independent)	Lien (KMT)	Chen (DPP)	Undecided & Others	Second Difference & (Total)	First Difference & (Total)	Measure 5	Modified Measure 5
Global/the Journalist	Jan 24	23	16.4	23.3	37.3		0.3 (46.3)	2.16	2.58
Poll Survey Organization	March 3	21.3	23.9	23.8	31		2.5 (45.1)	0.04	2.31
Extreme Case		10	37.54	12.46	40		2.46 (22.46)	0	7.27
Election Result	Mar 18	36.84	23.1	39.3	0.76	2.46 (76.14)			

The modified Measure 5 calculates the relative difference between the two margins, rather than the absolute difference. The effect of this modification resembles that of a proportional allocation of all supports other than the leading two candidates to the leading two candidates. This means that the second difference is modified as  $2.46/76.14 = 3.23\%$ , and the first differences are also modified as  $0.3/46.3 = 0.65\%$ ,  $2.5/45.1 = 5.54\%$ , and  $2.46/22.46$

=10.5% for Global/the Journalist, Poll Survey Organization, and extreme case, respectively.

As seen in Table 1, the modified Measure 5 has eliminated the unreasonable zero error calculation for the Extreme Case, but remains unexpected for the error calculation for Poll Survey Organization. According to Measure 5 or the modified Measure 5, it could be concluded that Poll Survey Organization had “good poll performance” because there was a small predicting error of margin for the top two winners. However, their data is obviously frail — it predicted that Lien would win, but he actually came in third in the election. Global/the Journalist’s poll, on the other hand, has predicted the exact right order as the three main candidates came in the poll, as well as revealed that the top two candidates had led by a safe margin with the third finishing just as it is in the actual election.

**Rules for Categorizing Polls:** Since Chen and Soong had such a clear cut (13.7%) lead over Lien, a mechanism is established to prevent the wrong measurement as that committed by Poll Survey Organization—“only the polls that picked correctly the two leading candidates have the opportunity to be judged as accurate.” The objective is to find a way to prevent the possible hidden error hoodwinked in Measure 5 by which the prediction of who came in third is wrong.

According to how the prediction on the third was made wrong, the polls are divided into the following three categories:

1. Predicted correctly the two leading candidates, meaning that the candidate who came in third was predicted correctly. Therefore, those polls with the order Soong-Chen-Lien or Chen-Soong-Lien are in category 1.
2. Predicted the correct winner, but predicted the second place candidate as the third. Therefore, those polls with the order Chen-Lien-Soong are in category 2.
3. The winner was predicted as the third or the reverse. Therefore, those polls with the order Lien-Chen-Soong, Lien-Soong-Chen or Soong-Lien-Chen are in category 3.

Among the three categories, category 1 is indeed superior to category 2 which is, in turn, superior to category 3. Obviously, the polls in category 3 have the worst performance since they predict the third as the winner or the reverse. Besides, the order Soong-Chen-Lien, as well as Chen-Soong-Lien, is in category 1, is superior to Chen-Lien-Soong which is in category 2. The logic is Chen-Soong-Lien erroneously foretold the order of a candidate who

actually came in third in the election and lagged behind both leading candidates by more than 13 percentage points. It would be a serious mistake. Therefore, the rule is simply defined as follows: polls sorted first category are judged more accurate than those in the second category, and so do the second category polls are judged more accurate than the third category polls. Furthermore, the polls in each category are ranked according to the error calculation from modified Measure 5.

## 2. Assessment

Table 2 presents 22 final poll results released by the 22 organizations, and the last column shows the poll category as we summarized previously. Only 9 of the 22 final polls predicted correctly the two leading candidates in this quasi-two-way election. Obviously, Taiwan's election polling had a terrible year in 2000. That motivates us to assess in more detail their poll performance individually.

For the convenience of comparing the results with Mitofsky's (1998: 230-249) and Traugott's (2001: 389-419), we calculate these 22 polls in Table 3 using Modified Measure 5, Measure 3 with proportional allocation (calculated after allocation of the undecided and the others), and Measure 5 without any allocation, (which were originally used in Mitofsky (1998: 230-249) and Traugott (2001: 389-419)), and the Chi-square statistic. Also shown are their corresponding error ranking that a rank of 1 was assigned to the poll with the smallest error for a given method. If more than one poll had the same error, they share their average rank.

Table 3 shows some unreasonable ranking results by Measure 5. For example, Chinese Society for Polls predicted the wrong winner as Lien, but was still judged as the rank 1 by Measure 5. In addition, there is no statistically significant difference between the average rank in different categories, while the average rank in category 1 is significantly smaller than those in category 2 or 3 by Measure 3 or Chi-square statistic. Furthermore, in Table 3, all the polls got similar rank values (maximum rank difference for polls is 3) for Measure 3 and Chi-square statistics. It is reasonable since both measures use the same information in their error calculation. On the contrary, the ranks induced by Measure 5 are much different from both measures. That is the reason why we modify the ranking procedure by Measure 5.

Table 4 compares the effect on the errors and rankings, which are assumed to be performance evaluating values, as calculated for Measures 3

**Table 2 Two Assessments of Accuracy in Taiwan's  
Final 2000 Presidential Election Polls**

Sponsoring/Polling Organization <sup>8</sup>	Date	Soong	Lien	Chen	Undecided & Others	Category
Global TV/the Journalist/CNA	1/17	25.3	16.6	25	33.1	1
Common Wealth	1/27	27.9	16.1	30	26	1
CTN (TV)	1/31	29.1	20.1	26.4	24.4	1
Central Police U.	2/19	23.8	25.2	21.3	29.7	3
SET TV/Gallup	2/29	23.2	15	22.7	39.1	1
KMT	2/29	24	26	23	27	3
China Academic Foundation	3/1	22.1	25.2	24.1	28.6	3
Power News	3/1	20.8	21.9	22.3	35	2
China Times	3/3	22	20	22	36	1
Shih Hsin U.	3/5	26.9	18.7	24	30.4	1
United News	3/5	26	27	22	25	3
Business Development Research Center	3/6	19.5	25.8	22.5	32.2	3
Decision & Public Survey Center	3/6	24.2	20.3	24.2	31.3	1
Mei-Yu Culture Foundation	3/6	21.7	24.1	22.1	32.1	3
Global TV/the Journalist/Focus	3/7	26.5	22.8	21.7	29	3
Chen's Campaign	3/7	25.1	22.4	25.8	26.7	1
Chinese Society for Polls	3/7	21.3	23.9	23.8	31	3
Competition Institution/Big Trendancy	3/7	22	24.5	21.3	32.2	3
Soong's Campaign	3/7	23.1	17.6	21.1	38.2	1
Lien's Campaign	3/7	30.9	34.6	29.4	5.1	3
ROC Women Institution	3/7	20.7	22.2	21.8	35.3	3
TVBS (TV)	3/7	24	25	26	25	2
Election Result	3/18	36.84	23.1	39.3	0.76	

<sup>8</sup> When there is a partnership between the actual polling organization and a sponsoring organization, they are separated by a slash. For instance, row 1 listed Global TV/ New News/CNA, where Global TV and New News sponsored the poll conducted by CNA, and the sponsors released the poll results published on 1/17. However in most cases, the polling and sponsoring organizations are the same.

**Table 3-1 Poll Accuracy Measurement, and rankings in Taiwan's Final 2000 Presidential Election Polls**

Sponsoring/Polling Organizations	Measure 3	Rank	Measure 5	Rank	Modified Measure 5	Rank	Chi- square	Rank	Category
Global/The Journalist/CNA	1.5	2.5	2.8	12	3.8	5	2.38	3	1
Common Wealth	1.0	1	0.4	2	0.4	1	1.28	1	1
CTN (TV)	3.1	4	5.2	19	8.1	8	11.22	4	1
Central Police U.	8.3	18	5.0	18	8.8	20	143.63	21	3
SET TV/Gallup	1.5	2.5	3.0	13	4.3	6	2.19	2	1
KMT	8.2	17	3.5	15	5.4	18	96.95	17	3
China Academic	8.0	15	0.5	4	1.1	13	83.38	14	3
Power News	6.9	12	1.0	6	0.2	10	57.50	11	2
China Times	5.3	9	2.5	10.5	3.2	3.5	37.02	9	1
Shih Hsin U.	3.4	7	5.4	20	8.9	9	12.76	5	1
United News	8.5	19	6.5	21	11.6	21	83.71	15	3
Business Development Research Center	9.8	22	0.5	4	3.9	16	148.73	22	3
Decision & Public Survey Center	4.2	6	2.5	10.5	3.2	3.5	29.11	6	1
Mei-Yu Culture Foundation	8.1	16	2.1	9	2.3	14.5	100.43	18	3
Global TV/The Journalist/Focus	6.0	10	7.3	22	13.2	22	54.94	10	3
Chen's Campaign	4.9	8	1.8	8	1.9	2	31.77	8	1
Chinese Society for Polls	7.5	14	0.0	1	2.3	14.5	73.45	13	3
Competition Institution/Big Trendency	8.6	20	3.2	14	4.8	17	102.58	19	3
Soong's Campaign	3.7	6	4.5	17	7.8	7	29.83	7	1
Lien's Campaign	8.8	21	4.0	16	5.7	19	108.81	20	3
Women Institution	7.3	13	1.4	7	0.6	12	85.82	16	3
TVBS (TV)	6.7	11	0.5	4	0.8	11	59.13	12	2
Election Result	36.84	23.1	39.3	0.76					

and 5, and modified Measure 5, and Chi-square statistics. Similiar results from what we have found in Table 3, the rankings produced by different methods, show that the results derived from Measure 3 and its Chi-square are extremely consistent ( $r=0.951$ ,  $p=0.000$ ). While the ranks produced by modified Measure 5 are more consistent with either measure 3 or the Chi-square statistics ( $r=0.651$  and  $0.558$ , respectively, each is statistically significant) than those produced by Measure 5 ( $r=0.011$  and  $-0.057$ , respectively, each is not statistically significant). That means the modified ranking results by modified Measure 5 are preferable than the original Measure 5.

### 3. Potential Sources of Error in Trial Heat Polls

As Lau (1994: 2-20) suggested, other non-sampling aspects of the survey research process affect fluctuating poll results. The polls may differ on eight independent variables, including sample size, percent undecided &



**Table 3-2 Poll Accuracy Measurement, and rankings in Taiwan's Final 2000 Presidential Election Polls**

Sponsoring/Polling Organizations	Measure 3	Rank	Measure 5	Rank	Modified Measure 5	Rank	Chi-square	Rank	Category
Global/The Journalist/CNA	1.5	2.5	2.8	12	3.8	5	2.38	3	1
Common Wealth	1.0	1	0.4	2	0.4	1	1.28	1	1
CTN (TV)	3.1	4	5.2	19	8.1	8	11.22	4	1
Central Police U.	8.3	18	5.0	18	8.8	20	143.63	21	3
SET TV/Gallup	1.5	2.5	3.0	13	4.3	6	2.19	2	1
KMT	8.2	17	3.5	15	5.4	18	96.95	17	3
China Academic	8.0	15	0.5	4	1.1	13	83.38	14	3
Power News	6.9	12	1.0	6	0.2	10	57.50	11	2
China Times	5.3	9	2.5	10.5	3.2	3.5	37.02	9	1
Shih Hsin U.	3.4	7	5.4	20	8.9	9	12.76	5	1
United News	8.5	19	6.5	21	11.6	21	83.71	15	3
Business Development Research Center	9.8	22	0.5	4	3.9	16	148.73	22	3
Decision & Public Survey Center	4.2	6	2.5	10.5	3.2	3.5	29.11	6	1
Mei-Yu Culture Foundation	8.1	16	2.1	9	2.3	14.5	100.43	18	3
Global TV/The Journalist/Focus	6.0	10	7.3	22	13.2	22	54.94	10	3
Chen's Campaign	4.9	8	1.8	8	1.9	2	31.77	8	1
Chinese Society for Polls	7.5	14	0.0	1	2.3	14.5	73.45	13	3
Competition Institution/Big Trendency	8.6	20	3.2	14	4.8	17	102.58	19	3
Soong's Campaign	3.7	6	4.5	17	7.8	7	29.83	7	1
Lien's Campaign	8.8	21	4.0	16	5.7	19	108.81	20	3
Women Institution	7.3	13	1.4	7	0.6	12	85.82	16	3
TVBS (TV)	6.7	11	0.5	4	0.8	11	59.13	12	2
Election Result	36.84	23.1	39.3	0.76					

**Table 4 Correlation Coefficient among the Ranks Derived from Different Measure**

Variables	Correlation (p-value)
Rank of Measure 3 vs. Rank of Chi-square	0.951 (0.000)
Rank of Measure 3 vs. Rank of Measure 5	0.011 (0.960)
Rank of Chi-square vs. Rank of Measure 5	-0.057 (0.801)
Rank of Measure 3 vs. Rank of Modified Measure 5	0.651 (0.001)
Rank of Chi-square vs. Rank of Modified Measure 5	0.558 (0.007)

others, number of days leading to election, types of sponsoring organizations (media or non-media), sponsorship (whether an organization sponsored a survey house to conduct the polling, or that they conducted the polls themselves), frequency of polls published by the same organization, field period (days in the field for polling data collection), and weekdays only (timing of

survey is conducted on weekdays).<sup>9</sup>

Using 68 pre-election polls published two months before the election, the correlation coefficient is calculated as an initial examination of the relationship between the eight independent variables and dependent variables of modified Measure 5. Due to missing data, the ordinal probit regression model excludes two independent variables, “field period” and “weekdays only,” which were not disclosed by many sponsoring/polling organizations. As seen in Table 5, only two of the eight correlations are statistically significant. It may be simply concluded as the following: Lower percentages of undecided & others, and closer the days to election are associated with smaller polling error. The conclusions meet our expectation or the results from Lau (1994: 2-20). Since we can conjecture that the closer the election, the more certain the voters should be about which candidate they are going to vote on the Election Day. Besides, the lower the undecided percentage also means the more certain the voters. In conclusion, Days to Election and Percent Undecided & Others are the two most important variables influencing poll accuracy.

In Table 6, the overestimation and underestimation of the votes for the three main candidates are calculated, and are regressed on the eight independent variables. In order to avoid possible huge noises from those biased polls with poor performance, we restricted our analysis to the 43 polls in the first category, that is, polls that predicted correctly the two leading candidates. On average, the prediction is an overestimated 7.19 percentage points for Lien, a significantly underestimated 1.88 percentage points for Soong, and significantly underestimated 5.13 percentage points for Chen.

Furthermore, we can examine the individual effect of each independent variable from the multiple regression models. It is interesting to find that if a poll reduces one percentage point of undecided, the underestimation for Soong is improved by 0.155 percentage points, and the overestimation for Lien would be improved by 0.169 percentage points. This observation points out that more undecided voters significantly tend to vote for Soong than expected. Besides, time until the Election Day has a significant contribution in reducing the amount of underestimation for Soong and so do the overes-

---

9 We also suspected whether “refusal rate” affected polling accuracy. However, we were unable to examine this factor because our analysis was constrained by the information released by the sponsoring/polling organizations, and the majority did not publish the refusal rate.

**Table 5 Correlation Coefficients and Ordinal Probit Regression for the Rank of Modified Measure 5<sup>10</sup>**

Variable	Range	Correlation Coefficient	Regression Coefficients in Ordinal Probit Model
Sample size (N)	858-1567	0.08 (p=0.608, n=43)	0.002 (p=0.617)
Percent undecided & others	21.7-43.2	0.39** (p=0.012, n=43)	0.554*** (p=0.003)
Days leading to election	11-70	0.278* (p=0.093, n=43)	0.234* (p=0.096)
Types of sponsoring organization	0: Media 1: Non-media	0.078 (p=0.617, n=43)	0.113 (p=0.558)
Sponsorship	0: sponsored only 1: did not sponsor other polling organization, conducted polls by self	0.013 (p=0.932, n=43)	3.562 (p=0.113)
Frequency of polls	1-9	0.02 (p=0.899, n=43)	0.127 (p=0.770)
Field Period (days)	1-27	-0.225 (p=0.2, n=34)	—
Weekdays only	0: Weekdays only 1: Include weekend	-0.018 (p=0.918, n=34)	—
Constant			-9.11* (p=0.143)

**Table 6 Regression of Over (or Under) Estimation of Vote for Each Candidate<sup>11</sup>**

Variable	Soong	Lien	Chen
Sample size	0.001 (p=0.517)	-0.001 (p=0.644)	0.0001 (p=0.946)
Percent undecided & others	0.155*** (p=0.005)	-0.169** (p=0.05)	0.013 (p=0.816)
Days leading to election	0.047*** (p=0.005)	-0.057** (p=0.027)	0.011 (p=0.519)
Types of sponsoring organization	-0.792 (p=0.393)	2.83** (p=0.048)	-2.048** (p=0.038)
Sponsorship	-0.524 (p=0.530)	1.415 (p=0.284)	-0.882 (p=0.313)
Frequency of polls	0.046 (p=0.751)	0.127 (p=0.578)	-0.179 (p=0.240)
Constant	-9.01*** (p=0.002)	12.877*** (p=0.005)	-3.818 (p=0.202)
R <sup>2</sup>	0.375	0.353	0.169
Adjusted R <sup>2</sup>	0.313	0.290	0.087
Mean	-1.88*** (p=0.000)	7.19*** (p=0.000)	-5.31*** (p=0.000)

10 Since the comparison of modified Measure 5 is only meaningful when the polls are in the same category of prediction, the analysis is based on the data of category = 1 (n = 43), \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

11 \* p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01

timation for Lien. Furthermore, comparing with non-media organizations, the overestimation for Lien and underestimation for Chen for media organizations are reduced by about 2.83 and 2.05 percentage points, respectively.

## V. Conclusion

Pre-election polling in Taiwan has become a popular and important tool for political parties, campaigns, the media, and the public. However, a closer look at the poll performance in Taiwan's 2000 election reveals questionable polling accuracy. While some polling organizations predicted the election results accurately, many failed to do a satisfactory job. The contributions of this paper are three-fold. First, it established criteria for suitable polling accuracy measurements in different election situations. Second, it assessed Taiwanese pre-election polling performance in 2000. Third, it examined the multitude of likely factors that influenced accuracy of poll results in this dynamic election.

We found that the choice of polling accuracy measurement should be taken into account with the election situation. It established criteria for suitable polling accuracy measurement in four possible election situations by categorizing the SSRC measures. In a two-way election or a quasi-two-way election like Taiwan's 2000, SSRC Measure 5 is superior to others. While Measure 5 is adopted in assessing poll performance, we found the terrible hidden error by those extreme "biased" polls. After modification, a fair way for ranking the poll performance by modified Measure 5 is shown.

Using the modified Measure 5 to detect the potential sources of polling error, our results showed that "percent undecided" and "days leading to election" are the two most influential factors affecting polling accuracy in Taiwan's 2000 presidential election. It is demonstrated that a poll tends to be more reliable if it is conducted closer to the Election Day and with lower percentage of undecided — it reduces the amount of underestimation for Soong and the overestimation for Lien. The analysis of over (under) estimation shows why and how these two factors have influenced polling accuracy and serves as an interesting evidence for the diffusion of the "dump-save" effect. Furthermore, our findings echoed Lau (1994: 2-20)'s empirical results that showed a significant correlation in the reverse direction where polls conducted closer to the Election Day should present smaller margin of error. Nevertheless, it is interesting to find that type of sponsoring organization is another significant factor impacting estimation, although it is not a signifi-

cant factor for polling accuracy. It shows that media organization tends not to overestimate Lien's vote but to underestimate Chen's vote.

Taiwan continues to present itself as an interesting model for measuring polling accuracy. In this paper, we had to implement proportional allocation for the high percentage of undecided voters before calculating the polling accuracy for each poll in 2000. However, in the 2004 presidential campaign, many polling organizations have developed their own prediction model for allocating undecided voters. Future research should continue to study the Taiwanese polling environment as it matures, including examining significant factors like "refusal rate" and establishing a historical reputation (quality) record for major sponsoring/polling organizations.

## References

- Alvarez, R. M and J. Nagler  
2000 "A New Approach for Modeling Strategic Voting in Multiparty Elections," *British Journal of Political Science* 30(1): 57-75.
- Blais, A. and R. Nadeau  
1996 "Measuring Strategic Voting: A Two-step Procedure," *Electoral Studies* 15(1): 39-52.
- Bowler, S. and D. J. Lanoue  
1992 "Strategic and Protest Voting for Third Parties: the Case of the Canadian NDP," *Western Political Quarterly* 45(2): 485-499.
- Buchanan, W.  
1986 "Election Predictions: An Empirical Assessment," *Public Opinion Quarterly* 50(2): 222-227.
- Burden, B. C.  
2005 "Minor Parties and Strategic Voting in Recent U.S. Presidential Elections," *Electoral Studies* 24(4): 603-618.
- CIA (United States Central Intelligence Agency)  
2007 "Taiwan, The World Fact Book," <http://www.cia.gov/library/publications/the-world-fact-book/print/tw.html> (September 7, 2007)
- Converse, P. E. and M. W. Traugott  
1986 "Assessing the Accuracy of Polls and Surveys," *Science* 234: 1094-1098.
- Crespi, I.  
1988 *Pre-election Polling: Sources of Accuracy and Error*. New York: Russell Sage.
- Crewe, I.  
1997 "The Opinion Polls: Confidence Restored?" *Parliamentary Affairs* 50(4): 569-585.

- Duverger, M.  
1954 *Political Parties: Their Organization and Activity in the Modern State North and North*. New York: John Wiley and Sons.
- Garand, J. C. and T. W. Parent  
1991 "Representation, Swing, and Bias in U.S. Presidential Elections, 1872-1988," *American Journal of Political Science* 35(4): 1011-1031.
- Huang, D. W. F.  
2004 "The Media, Opinion Polls, and the Election," <http://th.gio.gov.tw/p2000/p04.htm> (July 13, 2004)
- Lau, R. R.  
1994 "An Analysis of the Accuracy of 'Trail Heart' Polls During the 1992 Presidential Election," *Public Opinion Quarterly* 58(1): 2-20.
- Martin, E. A., M. W. Traugott, and C. Kennedy  
2005 "A Review and Proposal for a New Measure of Poll Accuracy," *Public Opinion Quarterly* 69(3): 342-369.
- Mitofsky, W. J.  
1998 "Was 1996 a Worse Year for Polls than 1948?" *Public Opinion Quarterly* 62(2): 230-249.
- Mosteller, F., P. J. McCarthy, E. S. Marks, and D. B. Truman  
1949 *The Pre-election Polls of 1948*. New York: Social Science Research Council.
- NCPP  
2007a "Final National Presidential Poll Results, 1936-2004," <http://www.ncpp.org/1936-2004.htm> (September 7, 2007)  
2007b "Presidential Poll Performance 2004 Error Calculator," [http://www.ncpp.org/err\\_calc.htm](http://www.ncpp.org/err_calc.htm) (September 7, 2007)
- Schoen, H.  
1999 "Split-ticket Voting in German Federal Elections, 1953-90: An Example of Sophisticated Balloting?" *Electoral Studies* 18(4): 473-496.
- Traugott, M. W.  
2001 "Assessing Poll Performance in the 2000 Campaign," *Public Opinion Quarterly* 65(3): 389-419.  
2005 "The Accuracy of the National Preelection Polls in the 2004 Presidential Election," *Public Opinion Quarterly* 69(5): 642-656.
- Wu, C. and D. Fell  
2001 "Taiwan's Party Primaries in Comparative Perspective," *Japanese Journal of Political Science* 2(1): 23-45.

# 準-兩位主要候選人選情下之 民調準確度評量： 2000 年台灣總統選舉之實證

蘇建州

世新大學傳播管理學系副教授

Mandy Sha

Research Survey Methodologist  
RTI International

## 摘 要

本研究主要是透過評量 2000 年台灣總統選舉民調準確度，探討如何在不同的選情採用不同的評量方法。從 2000 年台灣總統選舉結果顯示，該選舉的選情是介於兩位主要候選人與三位主要候選人的「準-兩位主要候選人」，然而是否不同的選情應該採用不同的民調準確度評量計算方法？本研究修改 Mitofsky and Traugott's 用以評量 1996、2000 與 2004 美國總統選舉（兩位主要候選人）民調表現的方法，提出一套比較合理適用於評量 2000 年台灣總統選舉民調的方法。本研究除了採用修正的方法分析選舉民調表現外，也進一步透過實證探討影響民調準確度差異的因子。

關鍵詞：民調準確度評量、民調表現、準-兩位主要候選人選情