



Forecasting the Election Results by Applying Pavia's Method

Nattapon Yamchim*, Pattrawadee Makmee & Kanok Pantong

College of Research Methodology and Cognitive Science, Burapha University, Chonburi, 20131 Thailand

Article info

Article history:

Received: 2 September 2019

Revised: 11 November 2019

Accepted: 29 November 2019

Keywords:

Forecast, Election, Opinion survey, Poll

Abstract

This research aims to forecast the election results by applying Pavia's method. In this paper, the information of opinion toward general election, which is reflected as one of behavioural science is applied, including using applied statistics to forecast the election results before announcing the election results. The process of data collection about people opinion was proceeded by survey related to election issues. In this survey, the sample was 3,600 electorates in the general election on 24th March 2019 from 30 electoral zones in Bangkok and the questionnaire about opinion of the general election was used as the tool for data collection. The applied statistics methods in this survey are percentage, Pavia's method analysis (Mean Absolute Percent Error: MAPE). The poll revealed that five parties received major scores, 22.69% for Pheu Thai, 21.94% for Democrat, 20.39% for Palang Pracharath and 16.69% for Future Forward Party. In terms of analysis by using Pavia's Method, the poll showed different results, 23.96% for Palang Pracharath, 22.45% for Future Forward Party, 21.25% for Pheu Thai Party and 19.12% for Democrat Party. When the poll results by using Pavia's method was compared with actual election, the percent of accuracy indicated at 82.28% or 17.12% of error.

Introduction

The public opinion survey or polling has been well accepted and received popularity in the United State of America (USA). This type of survey has been proceeded in order to predict or forecast the election results or interesting public issues.

The first forecasting simulation which was designed to forecast the result of the USA presidential election was used in Econometrics class, studying of the effect of economic activity on voting. The purpose of this study was to present the voting behavior and analytical pattern from the effect of economic activity on presidential voting (Fair, 1978).

In the case of USA presidential election, forecasting with regression model was used to forecast the state's voting results by using the past record of national polls and all information of all states in the USA. This forecast revealed that leading presidential candidate in any state in September before the election seemed to be elected on election day in November. The information of pre-election polls and post-variable was together applied for data analysis. This method increased the punctuality and accuracy of the USA presidential election forecasting. (Holbrook & Desart, 1999).

In 2010, Pavia had improved the accuracy of the forecasting of election results based on the results of the

polls from polling booths. The improvement consisted of three parts of the forecasting by using raw data from the direct surveys; Part 1 was the process of using bias checking to measure improvement of the use of Nonresponse Bias. Part 2 implemented an approximation to operate after the use of bias checking of Nonresponses to aid in the forecast. In addition, Part 3 was implemented after the process in the Part 2 was completed, resulting in the integration of different variables. When forecasting, regardless of the method of forecasting results, errors can occur. In 1948, the forecasting of the USA presidential election results showed an error between Dewey's and Truman's popularity. The forecast result revealed that Dewey would win but the election results with a score of more than 5% proved that Truman was elected president. This mistake caused a crisis of faith in the polls and polling agencies. The polls have been greatly improved (Wapor Exit Poll Committee, 2006; Jounes, 2008).; and yet in the 2016 election, an error in the USA presidential election forecasting was once again wrong when the results of the polls showed that Clinton would win, however, in fact, Trump was the winner.

In terms of Thailand, a general election poll was first proceeded in 1975 and the poll has been continuously performed by many polling agencies. When the parliament election took place on 3rd July 2011, many pre-election polls were surveyed by several agencies, Ramkhamhaeng Poll, ABAC Poll and Suan Dusit Poll all showed errors. For example, the Suan Dusit Poll's result predicted 162 seats of parliament were to be taken by Democrats, a small error proved the actual result of 165 seats. However, when considering the poll in each electoral zone, the forecasted number of members of parliament (MPs) from Suan Dusit Poll had a significant error. Three seats of MPs in Bangkok were forecasted to be 25 seats for Pheu Thai and 5 seats for Democrat. But, the actual election results showed dramatic figures, 27 seats for Democrats and only 9 seats for Pheu Thai. This figure gave a significant error around 44.45%.

As a result of this forecasting error, investigators determined to find out a forecasting method of election poll and improve forecasting results of polling by Pavia's Method. This forecasting method specifies how to improve accuracy of election forecasting from exit poll results and also uses bias testing for improving personal bias when answering the question. However, personal information or background of the electorate was not applied in this forecasting improvement. Also, Trangucci, Ali, Gelman, & Rivers (2018) explains voting

pattern in 2016, analysis of pre-election poll in 2012 and 2016. This research shows difference between voting result in each electorate group. Gender and education level play a significant role in the different voting decision. Moreover, personal and background information of electorate was applied together with Pavia's method adjustment in order to improve the accuracy of election forecasting, when compared with the actual election result.

Objective

To forecast the results of member of parliament election in 2019 by applying Pavia's method.

Conceptual framework

Election forecasting by using opinion poll and statistical analysis creates forecasting efficiency and reduces forecasting error. The Pavia's method is also recognised as one of the developed methods which suggests the way to improve forecast accuracy

$${}_{2ST}\hat{Z}_j = \sum_{b=1}^s \bar{\omega}_b \hat{p}_{jb}$$

by using the exit poll and past election results in election forecasting. Also, application of regression model without using other related variables, electorate, candidate and economic information is used. These variables may have an affect to election forecasting and finally may have the ability for eliminating the errors. Personal information of the electorate was not used in this analysis. Therefore, the researchers decided to adjust P-Value by using logistic regression model that includes the background information of the electorate.

From the literature review of the electorate's decision factor, personal background and individual information all have important effects on the decision-making such as gender, age, occupation, educational level, income and racial origin (Trangucci, Ali, Gelman, & Rivers, 2018). In this paper, individual information provided by Suan Dusit Poll, gender, age, occupation, educational level, income without racial origin are used.

The conceptual framework of the study to develop forecasting election result is explained in figure 1 as shown below:

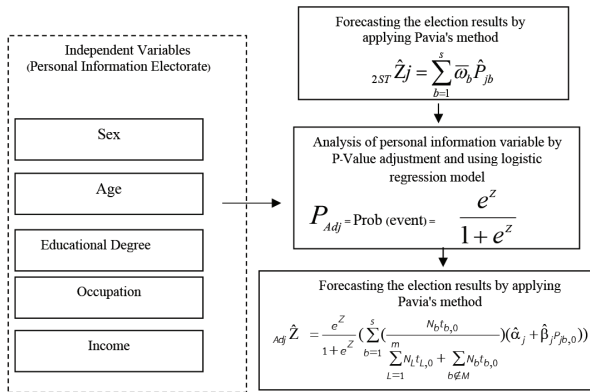


Figure 1 Conceptual framework

Research methodology

1. Population and sample

The population: electorate for members of parliament election in 2019, 51,239,638 people

The population in this survey are the electorate who are voting for members of parliament election in Bangkok, 30 electoral zones, 5,701,394 electorate (official announcement from Election Commission of Thailand).

Sample: electorate for members of parliament election in 2019. this is the primary information from an opinion poll of members of parliament election in 2019 in 30 electoral zones of Bangkok. 3,600 samples were received by random sampling with probability sampling. The random sampling process was performed by Multi Stage Random Sampling Method.

Step 1: Cluster Sampling by arranging area in Bangkok to be 30 electoral zones

Step 2: Simple Random Sampling 50% of all residential districts in each electoral zone

Step 3: Simple Random Sampling electorate in the residential districts, which is chosen in step 2

2. Instrument construction and data collection

The instrument in this research is a questionnaire about the opinion on members of parliament election in 2019

This questionnaire is divided into two sections as follows:

Section 1: Check list about general demographic information, gender, age, occupation, educational level and income

Section 2: Check list about the opinion on rights for voting members of parliament (MPs)

Construction and Efficiency of The Instrument

The questionnaire in this survey is created and designed by several procedures as follows:

1. Setting up the questionnaire objective for constructing the questionnaire about members of parliament election

2. Conducting literature reviews

3. Specifying operational specific terms about the opinion on members of parliament election to be guideline of the questionnaire construction

4. Constructing the questionnaire following operational specific terms

5. Submitting the test created for 5 experts to check the accuracy. Then calculate the internal consistency index by the IOC and select the IOC value greater than 0.5 and adjust it to be appropriate and correct along with the recommendations of experts,

Due to this questionnaire included the personal opinion about election, reliability testing is not performed after receiving recommendations from experts

6. Producing copies of the questionnaire for further data collection.

3. Data collection

The data collection process was performed by researchers as following:

1. Planning data collection by scoping electoral zone

2. Preparing the sufficient questionnaire for population in this survey

3. Explaining to populations about purpose of data collection

4. Evaluating answered questionnaire by statistical method and testing hypothesis.

4. Forecasting the election results by applying Pavia's method

After receiving information about the opinion of Bangkok governor election in 2013, Pavia's method is used to edit forecasting together with applying background information of the electorate (${}_{Adj} \hat{Z}_j$)

$${}_{Adj} \hat{Z}_j = \sum_{b=1}^s \bar{\omega}_b \hat{P}_{jb} P_{Adj}$$

$$\bar{\omega}_b = \frac{N_b t_{b,0}}{\sum_{l=1}^m N_l t_{l,0} + \sum_{b \in M} N_b t_{b,0}}$$

$$\hat{P}_{jb} = \hat{\alpha}_j + \hat{\beta}_j P_{jb,0}$$

$$P_{Adj} = \frac{e^z}{1 + e^z}$$

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_p X_p$$

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$ = approximated coefficient from data

$X_1, X_2, \dots, X_p$ = variables of background information

$P_{jb,0}$ = registration ratio of previous election

N = the number electorate in electoral zone

N_L = the number of voter in electoral zone L

N_b = the number of voter in electoral zone B

M = set of electoral zone in sample area

L = survey area

b = electoral zone

s = all electoral zone

m = the number of sample area

5. Comparison between forecasting result and actual result by using total absolute

When forecasted results are applied to the Pavia's method, the forecasting error will be calculated by using total absolute between forecasting results and actual results. The forecasting error by using the background information of the electorate must be less than the previous forecasting results.

$$E = \sum_{j=1}^k |z_j - \hat{z}_j|$$

k = the number of forecasted party

z_j = election result of j party

$\hat{z}_j$ = the forecasting result of j party

6. Data Analysis and Statistics

Percentage, logistic regression model and error measurement are used to forecast the result of members of parliament election in 2019 by applying Pavia's method.

Result

Comparison of forecasting error by using the Pavia's method and actual results of the members of parliament election in 2019. This election was held on

24th March 2019 and the survey was proceeded before knowing the exit -poll. In this survey, 3,600 samples are used from 30 electoral zones in Bangkok. After that, the information was evaluated and analyzed by using the Pavia's method and compared with actual election results. The information can be summarised as follows:

From 3,600 samples, 50.83% of the population are female and 49.17% is male. For ageing, 3.22 % are in the range of 18-27 years, 12.58 % for the range of 28 – 37 years, 22.51 % for the range of 38-47 years, 28.22% for the range of 48-57 years and 33.47% for the range of 57 years and over.

For information of educational degree, 71.97% of the population hold the undergraduate degree, 25.69 % received a bachelor's degree and 2.349% received a postgraduate degree. For occupational information of the population, 5.78% are students, 11.31% work as government officers, 13.78% work as full-time employees, 33.25% work as business owner, and 35.88% are self-employed.

For income information of the population, 8.69% receive less than 5,000 Baht, 21.75% receive 5,000-10,000 Baht, 36.76% receive 10,001-20,000 Baht, 21.94 % receive 20,001-30,000 Baht and 10.86% receive more than 30,000 Baht.

The opinion poll before the election date identified that 17 parties will be elected. Pheu Thai received the highest voting at 22.69%, followed by 21.95% for Democrats, 20.39% for Palang Pracharath, 20.39% for Future Forward, 16.70% for Thai Liberal, 7.58% for Bhumjaithai, 4% for Puea Chat, 1.83% for Action Coalition for Thailand, 1.28% for Charthai Pattana, 1% for Thai Local Power, 0.69% for Chartpattana and 0.61% for another six parties.

From the result of the members of parliament election in 2019, in 30 electoral zones of Bangkok, there were 3,101,010 voters. Future Forward Party received the highest score and accounted for 25.936%, followed by Palang Pracharath at 25.537%, Pheu Thai at 19.5% and for Democrats 15.312%. Numbers should be rounded and match the chart below

From the results of the members of parliament election in 2019, 49 parties were elected with different score and three parties were elected form a total of 30 available seats for members of parliament in Bangkok, 12 seats for Palang Pracharath, 9 seats for Future Forward and 9 seats for Pheu Thai. The percentage of election score for each party is 25.94% for Future Forward, 25.54% for Palang Pracharath, 19.5% for Pheu Thai,

15.31% for Democrats, 4.52% for New Economics (NEP), 3.09% for Thai Liberal, 1.41% for Bhumjaithai, 1% for Action Coalition for Thailand and less than 1% for the other 41 parties.

Table 1 Comparison between survey result and actual election result of member of parliament election in 2019.

Party	Percentage of actual election result	Percentage of Survey Result	Percentage of Difference
Future Forward	25.94	16.69	9.24
Palang Pracharath	25.54	20.39	5.15
Pheu Thai	19.504	22.69	3.19
Democrats	15.314	21.94	6.63
New Economics (NEP)	4.524	0.08	4.44
Thai Liberal Party	3.09	7.58	4.50
Bhumjaithai	1.41	4.00	2.60
Action Coalition for Thailand	1.00	1.28	0.28
Puea Chat	0.90	1.83	0.93
Thai Local Power	0.37	0.69	0.33
Others	2.43	2.81	0.37

From Table 1, The comparison between survey results and actual election results of members of parliament election in 2019, the figure indicates error of pre-election poll and actual election results. The error value is mostly founded in Future Forward Party for 9.24% followed by Democrats, Palang Pracharath, Thai Liberal Party, New Economics (NEP) and Pheu Thai at 6.63%, 5.15%, 4.50%, 4.44% and 3.19%, respectively.

Forecasting the election results by applying Pavia's method This forecasting contains some information, registration ratio of previous election, number of electorate in electoral zones, number of voters in electoral zones, set of electoral zones in sample area, survey area, electoral zone, all electoral zones and number of sample areas. This information is described in table 2.

Table 2 Comparison of forecasting error by using Pavia's Method and actual results of the members of parliament election in 2019

Party	Pavia's method application	election result	Survey result
Future Forward	22.45	25.94	16.69
Palang Pracharath	23.96	25.54	20.39
Pheu Thai	21.25	19.50	22.69
Democrats	19.12	15.31	21.94
New Economics (NEP)	1.12	4.52	0.08
Thai Liberal Party	5.21	3.09	7.58
Bhumjaithai	2.00	1.41	4.00
Action Coalition for Thailand	0.82	1.00	1.28
Puea Chat	0.84	0.91	1.83
Thai Local Power	0.21	0.37	0.70
Others	3.02	2.43	2.81

From table 2, the figure shows error of forecasting results at 23.96% for Palang Pracharath (election result at 25.54%), 22.45% for Future Forward (election result at 25.94%), 21.25% for Pheu Thai (election result at 19.50%), 19.12% for Democrats (election result at 15.31%), 1.12% for New Economics (election result at 4.52%), 5.21% for Thai Liberal Party (election result at 3.09%), 2% for Bhumjaithai, (election result at 1.41%), 0.82% for Action Coalition for Thailand (election result at 1%), 0.84 % for Puea Chat (election result at 0.91%), 0.21% for Thai Local Power (election result at 0.37%) and 3.02% for other parties (election result at 2.43%).

Table 3 The percentage of discrepancies between the 2019 members of parliament election results and the forecasting of elections based on Pavia's Methods and survey results.

Party	Percentage of discrepancies	
	2019 members of parliament election results and the forecasting of elections based on Pavia's methods	2019 members of parliament election results and survey results.
Future Forward	3.49	9.25
Palang Pracharath	1.58	5.15
Pheu Thai	1.75	3.19
Democrats	3.81	6.63
New Economics (NEP)	3.40	4.44
Thai Liberal Party	2.12	4.49
Bhumjaithai	0.59	2.59
Action Coalition for Thailand	0.18	0.28
Puea Chat	0.06	0.92
Thai Local Power	0.16	0.33
Others	0.59	0.38

From table 3, it was found that three parties were elected from 30 electoral zones in Bangkok, 12 zones taken by Palang Pracharath, 9 zones taken by Future Forward and 9 zones taken by Future Forward. Also, an error is showed at 1.58%, 1.75% and 1.58% for Future Forward, Phue Thai and Palang Pracharath, respectively.

The least error was found for Puea Chat at 0.06% and the highest error found was for Democrats at 3.81%.

The forecasting error by using Pavia's method and actual results of the members of parliament election in 2019 is described in the bar graph below to show errors for each forecasting method.

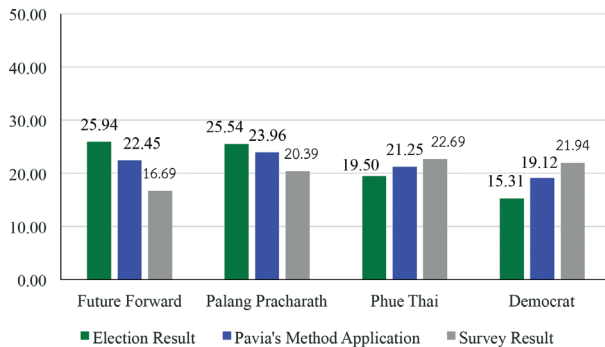


Figure 2 Comparison of the election results between applying Pavia's Method and actual election results

From figure 2, the figure shows that applying Pavia's method for election result forecasting has a small difference with the actual election results. This means applying Pavia's method can decrease forecasting error and create more accuracy of the forecasting.

Table 4 Forecasting error by applying Pavia's Method in the election of members of parliament in 2019

	Applying Pavia's Method with actual election results	Survey and Actual Election Results	Percentage of Decreased Error
Forecasting Error	17.72	37.65	52.94

From table 4, an error before applying Pavia's method, compared with actual election results of member of parliament in 2019 is at 37.65 %. However, after applying Pavia's method, an error is indicated at only 17.72%. This means accuracy of forecasting is significantly increased to be 82.28%.

Discussion

Forecasting the election results by applying Pavia's method shows a smaller error than forecasting the election result by polling before election, according to hypothesis no.3. The result of this research also proves that survey information before the election has error at 37.65%, compared with actual election result in 2019. However, using Pavia's method to forecast election results has smaller error at 17.72%. This means Pavia's method can decrease 19.93% of error, accounting for 52.94%. Nevertheless, error from using Pavia's method is still higher than expectation value from hypothesis, less than 10% of error.

The decreased error in the forecasting corresponds

with the study of Armstrong (2006), the study was conducted for more than 25 years, explaining forecasting by using evidence, reveals new seven methods for forecasting which are categorised into three groups;

1. The method being capable to apply to all types of information, combination of forecasting method which can reduce error at 12%, including Delphi method which can improve comparison accuracy, 19 of 24 subjects, accounted for 79%.

2. Cross-Sectional Data consists of Causal Models which can decrease 10% of error, Judgmental Bootstrapping, which is able to decrease 6% of error and Structured Judgment, that is unable to evaluate forecasting error.

3. Application of Time Series Data, Damped Trend, that is able to decrease 5% of error while Causal Models improves accuracy for 3/4 of intermediate and the long- term forecasting.

Pavia (2005) explored forecasting without random sampling including case study of the night before election in 1990-1999, the study presents forecasting procedures of the night before the election, that includes information exchange of political parties or candidates taken place and this showed correspondence with actual election results from each polling station, especially forecasting about the last ratio of election results for each party by using past and upcoming opinion survey from each polling station. The Pavia's method is still continually edited and improved and it has high flexibility and efficiency which can apply to forecast the general election. Finally, researchers concluded that problems or obstacles can normally happen for any forecasting. Therefore, in order to receive accurate forecasting results, the Pavia's method has to be applied with other methods of problem solving in random sampling, such as, random sampling of the poll, insufficient dispersion of sample.

In keeping with the study by Pavia (2010) that improved the accuracy of forecasting. The study improved accuracy of forecasting results from polling results of polling booths by using the survey data on SigmaDos. The forecasting process introduced an 11-part of prediction by comparing it with SigmaDos. In the first three parts of the forecast, the use of raw data from direct surveys and the use of bias checking was implemented to measure the improvement of the Nonresponse Bias. Part 2 presented an approximation to be made after the use of bias correction, unanswered questions which helps to determine whether it is valuable. Moreover, in Section 3, the estimation of the other four estimates which are

obtained from Part 2, with the integration of different variables, resulting in improvements that can improve forecasting to be the actual election results and reduce discrepancies.

Forecasting studies that use individual status variables that effect the forecasts of error predictions are Trangucci, Ali, Gelman, & Rivers (2018). They studied the voting pattern in 2016 which were surveyed by using Multilevel Regression and Poststratification: MRP in the pre-election of the year 2012 and 2016. The differences among the population who voted in the 2012 and 2016 elections were to divide the data by demographic and status information. It was found that (1) the gender gap had increased. In 2016, most government data relating to the study showed a U-Shaped curve displaying the larger gender gap of lower and higher education levels. (2) Younger white electorates who were not well educated gave more support to Donald Trump compared to younger electorates with higher education. (3) There were more women supporting Hillary Clinton than men, in addition, women who were young and highly educated support Hillary Clinton. (4) Elderly men with little education supported Donald Trump. (5) Color skin electorates overpoweringly supported Hillary Clinton. (6) The gap between electorates educated in colleges and those who are not educated in the college had approximately 10% of devotion of Hillary Clinton. The study of forecasting error also discovers similarity with Hibbs (2000), explaining "Bread and Peace" Model which indicates two explanatory variables), economic prosperity (Bread) and Peace. This Model gives well-explanation of USA presidential election from 1952-1992 but suggests incorrect forecasting of the election in 1996 and 2000. "Bread and Peace" Model specifies that a perfect forecasting model for USA presidential election does not truly exist and a powerful model must be completed by model improvement, all the time. These models play significant role as "Mind Reader" of the people. If primary behavior of electorate is changed, the previous or current forecasting model can give incorrect result due to structural problem of the model.

In Conclusion, from this research it can be concluded that forecasting the election results by applying Pavia's method by using background information of the electorate provides smaller forecasting error. This forecasting method has more accuracy than previous forecasting method which corresponds with hypothesis of the study.

Suggestions

1. For forecasting the election results, other related and external factors such as political situations, economic conditions, number of candidates, election system et al., have to be concerned in election forecasting.
2. Application and development of forecasting model has to consider the specific context of the election in each area.
3. Conditions and agreement have to be considered when choosing applied statistics for election forecasting.

References

- Armstrong, J. S. (2006). Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting*, 22(3), 583-598.
- Hibbs, D. A. (2000). Bread and Peace Voting in US Presidential Elections. *Public Choice*, 104(1), 149-180
- Fair, R. C. (1978). The Effect of Economic Events on Votes for President. *The review of Economics and Statistics*, 60(2), 159-173.
- Holbrook, T. M., & DeSart, J. A. (1999). Using state polls to forecast presidential election outcomes in the American states. *International Journal of Forecasting*, 15(2), 137-142.
- Pavia, J. M. (2005). Forecasts from nonrandom samples: The election night case. *Journal of the American Statistical Association*, 100(472), 1113-1122.
- Pavia, J. M. (2010). Improving predictive accuracy of exit polls. *International Journal of Forecasting*, 26(1), 68-81.
- Jones Jr, R. J. (2008). The state of presidential election forecasting: The 2004 experience. *International Journal of Forecasting*, 24(2), 310-321.
- Trangucci, R., Ali, I., Gelman, A., & Rivers, D. (2018). *Voting patterns in 2016: Exploration using multilevel regression and poststratification (MRP) on pre-election polls*. Retrieved November 8, 2019, from <https://arxiv.org/pdf/1802.00842.pdf>
- Wapor Exit Poll Committee. (2006). *WAPOR Guidelines for Exit Polls and Election Forecasts*. Retrieved November 8, 2019, from <https://wapor.org/wp-content/uploads/WAPOR-Guidelines-for-Exit-Polls-and-Election-Forecasts.pdf>