

USING STATISTICS AS A TOOL IN POLITICAL RESEARCH. THE CASE OF ELECTORAL BEHAVIOR

Theodore Chadjipadelis
Aristotle University
Greece

In this paper an extensive reference to the problems of using statistical techniques in the political sciences is taken place. Statistics today, in the era of information outbreak can be generally defined as the "Science that is concerned with the gathering, evaluation and processing of information". Society's demand for qualitatively controlled information absolved from "noises" which intentionally or not are included, is especially obvious in a returning, from time to time discussion, about the control of the public measurements related to the Mass Media audience, political parties and persons, educational parameters, economical and social indexes and factors, etc. By following and commenting the phases of observation of electoral behavior, we will refer to some problems.

INTRODUCTION

Prior to electoral processes the civilians are bombarded with information, graphs, tables and findings of opinion polls which are referred in indexes, figures such as the expected power of parties participating in elections, the publicity of persons, evaluative classification of issues and politicians. After the elections, the results are explained and analyzed by using advanced statistical techniques aiming to support the pre-electoral statements and attitudes of political parties. While publishing the results, many analyzers and journalists do not understand that the values that are counted for the figures (usually percentages or flows) are not accurate values but estimations. They do not understand that a confidence interval means that the estimated parameter e.g. the percentage of a political party is found with exactly the same probability between the upper and lower limit. They believe that this parameter is found nearer to the center of the confidence interval.

All political parties are interested in the analysis of behavior of subpopulations of people. However, if the sample can be designed so that to represent the general population, this does not mean that each part of it represents the subpopulation from which it happened to come from. The supporters of some party found in the sample are not a representative sample of the total supporters of the party, and the youth between 18 and 25 found in the sample are not necessarily a representative sample of all the youth, etc. The use of methods of analyzing two-way tables (such as correspondence analysis) can lead to errors and misunderstandings.

Statistics is not a total of rules and recipes for the analysis of data. It is not exhausted in the use of complicated computerized programs and nice graphs. It requires a good knowledge of the observed phenomenon, the organization of observation, good knowledge of gathered data, and description and examination of hypotheses for the parameters of the analyzed phenomenon. In the level of Higher Education Statistics is taught as an obligatory subject in all university departments with a main goal the students' comprehension for the use of statistical techniques orientated to the specific field of knowledge. Although students of so-called "theoretical" departments consider that they do not have a good relation with numbers, mathematics, statistics and computers, the need of experimental research leads them to the need of understanding and using not only of simple descriptive methods, but especially advanced statistical techniques which demand mainly "mathematical thought", that is the ability of constructing, using and explaining abstract models from daily examples. By following and commenting the phases of observation of electoral behavior we will refer to some problems.

ELECTORAL BEHAVIOR

The methods of prediction and observation of electoral (and widely the political) behavior of the electorate have a long history of application and use in the countries of Western Europe and the U.S.A. and they decisively depend on the specification of each country (political history, electoral system, electoral framework, etc.), we can be divided in two sub-categories:

Opinion polls and PANEL researches. In the first category there is an effort to get results about the behavior of the total of voters often with repetitive polls on the same or similar sample of the electorate. As it is in every opinion poll, the initial problem is the selection of the sample, which should be representative (of the population) so that the drawn results could be generalized by the use of the proper statistical techniques.

SAMPLING, CHOICE OF SAMPLE

In the formation of the sample some general factors (such as sex, age, social and financial status), which influence electoral behavior should also be considered, as well as the previous behavior of the electorate (results of previous elections, behavior of the sample in previous elections, etc). In Greece the formation of mixed areas of residence (which are not necessary defined by specific social groups) makes us to see skeptically geographical methods of selecting a sample, and we rather pay more attention to criteria of electoral behavior for the selection of the sample. Of course, when the electoral behavior of the electorate is homogenous, (they all behave in almost the same way), whether we follow a geographical selection of sample or not, the result of the poll will be the same. Accordingly, our results cannot be especially analytical because there is no identification of the electorate with the resident-voters of the area. The formation of sample in all cases should be based on criteria of previous electoral behavior in specific areas based on the analysis, with advanced statistical techniques, of the results of the previous electoral data, in the level of voting areas.

The analysis should be based in the estimation of the flows from state to state. Flow is the central size of political changes, according to the bibliography, and the specific sizes with the analysis of comparative data of the sample without the influence of weightings, which alter the first characteristics. It is known that the size of the selected sample affects the basic results (e.g. for the percentages of the political parties we calculate 95% confidence interval with 2-4 percentage units of width). This means that with a probability of 95% the estimated parameters (e.g. percentages) will be within the limits of the confidence interval. If these values move off the limits of the confidence interval, irrelevant to the numerical difference, they are proved as wrong.

QUESTIONNAIRE, A TOOL OF MEASUREMENT

Questionnaires are often designed as simple lists of questions to be answered and not as cohesive total-tools. Especially when a multivariate analysis technique is going to be used (such as e.g. FACTOR or DISCRIMINANT) for the drawing of results, the above can lead in many misunderstandings and problems.

EXIT POLLS

This method consists of the selection of a “*representative*” sample on the elections day and of the results about the final outcome by the sample. At this point we should distinguish the method of exit poll from the methods of predictions based on real results from the ballot box. As it becomes obvious, the exit poll carries all the characteristics of an opinion research with the decisive “*advantage*” that the percentage of indecisive voters is dramatically reduced (actually it is limited to those refusing expression of opinion). Also in this case the sample has to be formed using the criteria of electoral behavior rather geographical ones. Also we should rather observe the voters’ flows than the absolute percentages. Furthermore we face some difficulties such as: a part of the sample does not have a registered electoral behavior (new voters) and the questionnaire must be especially simple, clear and small since it should be completed by the person asked – if this is possible – in the absence of the interviewer. In its first application in Greece this method had to face the special phenomenon of “non-resident” voters, i.e. a part of the electorate voting in their place of residence and not in the place of registration as voters. This fact created some difficulties in the prediction of flows of voters because of the increasing of that part of the sample, which did not have a previous recorded behavior. It is estimated that an important part of the “no-resident” voters voted differently in their place of residence than it would vote in the place of registration. Exit polls give at the same time valuable data for the extract of a complete picture especially about the flows between political attitudes of elections that cannot be imprinted on the election results.

Table 1
Results of Parliamentary Elections in 1993 and 1996

Political Party	Parliamentary Elections 1993	Parliamentary Elections 1996
Void	2,33	2,84
PASOK (<i>Panhellenic Socialist Party</i>)	46,88	41,52
ND (<i>New Democracy</i>) <i>right wing party</i>	39,30	38,14
POLAN (<i>Political Spring</i>) <i>right wing party</i>	4,87	2,94
KKE (<i>Communist Party of Greece</i>)	4,54	5,60
SYN (<i>United Left Party</i>)	2,94	5,10
DIKKI (<i>Democratic Socialist Movement</i>)	*	4,43
Other	1,47	2,94

*Members and supporters of PASOK formed DIKKI in 1994

The comparative observation of results of the two recent parliamentary elections (Table 1) it is not enough to estimate these flows. For example, observing that the sum of percentages of DIKKI and PASOK in the Parliamentary elections in 1996 equals to the percentage of PASOK in the previous elections we can assume that 10% of the PASOK 1993 voters have chosen DIKKI, as it was already estimated by polls prior to the 1996 elections (Table 2). These results contradict the percentage of flow about 6% estimated by exit polls.

Table 2
Estimated Flow of Voters from PASOK to DIKKI

ICAP (Eleftherotypia Newspaper 12/9)	10,9%
MRB (Eleftherotypia Newspaper 9/9)	11,0%
RI (NEA Newspaper 12/9)	8,6%
ALKO (Free Press 12/9)	9,7%

In recent parliamentary elections exit polls were based on selections of giant samples (10000 – 30000) citizens and the conducting of preliminary opinion polls of a wide range for the estimation of random local differentiation. To what extent the selection of *different* giant samples is generally approved and useful for the aim of fast information is still for discussion. The question of “*why should we pay 60 –100 million drachmas to have the result 3-4 hours earlier*” makes sense. In contradiction, a small compact sample cannot certainly give the slight differences of “*local vote*” i.e. the specific flows in each voting region. That remains to be estimated with other methods from the final results. Anyway due to the design of the sample we get the general characteristics of flows. The estimation of the average gives accurately the general characteristics of election results in central figures but it is not easy to give the exact result, as it is shown from the following comparative table (Table 3).

Table 3
Exit Polls 1996 (Greece)

Parties	Organization/company conducted the poll						Election RESULT
	PRC	OPINION BVA	ALCO	KAPA RESEARCH	EURIN COM	A.U.TH	
PA.SO.K.	42,1	42	41,7	41,8	41,7	41,64	41,51
N.D.	38,9	38	38,5	38,9	38,8	38,32	38,16
POL.AN.	3,3	3	3,3	3,3	3,2	3,11	2,94
K.K.E.	4,9	5,5	4,9	5,2	5,4	4,80	5,60
SYN	4,5	4,5	4,9	4,6	4,7	5,50	5,10
DI.K.KI.	4,0	4	3,9	4,1	3,9	4,61	4,43
OTHER	2,4	3	2,8	2,1	2,3	2,03	2,84

The estimation of results for the country was conducted by the author (*A.U.TH*) in a representative sample (of 1161 citizens of the area A' in Thessaloniki), with a confidence interval $\pm 2,5$ for the two first parties and $\pm 1,5$ for the rest and around the above central values, though other estimations were giving an interval from ± 1 to $\pm 0,5$ around the central values.

As it was mentioned above, the central finding of exit polls are the table of flows between attitudes in continual elections especially concerning the total of the country. The table of flows, computed for each voting area, gives results about the structure of the supporters of the political parties, considering the local features, and the general table of flows gives results about the flows throughout the country. These figures (percentages of flows from attitude to attitude) are also estimated central values and for this reason they do not have an accurate numerical significance. Nevertheless, it is generally an arbitrary result that the way voters behave in the total of the country they also behave in each voting area. In order to estimate whether these two "pictures" are identical, (i.e. whether the flows in each voting region are similar to the general flows throughout the country), the final results should be used, in order to get estimations of the marginal flows in voting geographical units. The following tables represent the flows in a selected sample of 480 voters in the A' region of Thessaloniki. The flow in each cell on the table represents the percentage of voters of the political party, which corresponds with the line on the table the party voted corresponding on the column of the table. An additional indication about the difference of flows comes out from the comparative observation of the tables of flows which were estimated for the country and the A' region of Thessaloniki by the exit poll conducted in Thessaloniki (Tables 4 & 5).

Table 4

Flows for the Country from Lines (B93) to Columns (B96)

	PASOK	ND	POL. AN	KKE	SYN	DIKKI	OTHER	VOID	N.A
PA.SO.K.	81,47	2,46	1,12	1,56	2,01	6,25	0,67	2,01	2,46
N.D.	3,58	89,53	1,65	0,28	0,55	1,65	0,83	0,28	1,65
POL.AN	13,51	10,81	43,24	0,00	5,41	18,92	2,70	5,41	0,00
K.K.E.	8,93	1,79	0,00	85,71	3,57	0,00	0,00	0,00	0,00
SYN	7,25	4,35	0,00	4,35	78,26	1,45	1,45	0,00	2,90
NO VOTE	7,14	7,14	0,00	0,00	7,14	7,14	57,14	7,14	7,14
OTHER	0,00	50,0	0,00	0,00	50,0	0,00	0,00	0,00	0,00
VOID	20,0	13,33	0,00	0,00	13,33	0,00	6,67	40,0	6,67
N.A.	20,0	2,86	0,00	2,86	2,86	4,29	1,43	7,14	58,5
NEW VOTER	34,72	30,56	0,00	2,78	11,11	2,78	4,17	1,39	12,5

Table 5

Flows for the A' Region of Thessaloniki from the Lines (B93) to Columns (B96)

	PASOK	ND	POL AN	KKE	SYN	DIKKI	OTHER	VOID	N.A.
PA.SO.K.	84,07	2,01	0,37	1,47	2,93	5,49	1,83	0,92	0,92
N.D.	2,92	88,33	1,04	0,63	1,25	1,25	1,67	1,04	1,88
POL.AN	37,14	11,43	25,71	2,86	2,86	5,71	8,57	0,00	5,71
K.K.E.	2,86	0,00	0,00	90,0	2,86	1,43	0,00	1,43	1,43
SYN	2,63	1,32	0,00	3,95	84,21	3,95	0,00	0,00	3,95
NO VOTE	7,32	9,76	0,00	0,00	2,44	21,95	53,66	0,00	4,88
OTHER	10,0	0,00	0,00	0,00	30,0	10,0	30,0	10,0	10,0
VOID	4,65	4,65	4,65	2,33	18,6	4,65	6,98	53,5	0,00
N.A.	15,28	8,33	2,78	1,39	4,17	8,33	2,78	4,17	52,8
NEW VOTER	35,15	23,08	4,40	3,30	8,79	4,40	8,79	3,30	8,79

The picture of flows that comes out for the A' region of Thessaloniki is perceptibly different from the picture of the total country. We indicate that the numbers in each table are estimations from sample, i.e. central values (not absolute numbers). However, it is obvious that the flow from POLAN to PASOK in Thessaloniki is perceptibly larger than the one in the general table of the country which is explained from the influence by members of the Parliament of PASOK (the Government' Party) towards the "audience" of POLAN, as it was recorded in the previous polls from 93 till 96 in Thessaloniki. The flows in the specific voting region from KKE and SYN towards PASOK are perceptibly smaller than in the country, something that is explained by the relevant slight interaction of SYN and KKE with PASOK, and from the fact that this region is mainly an urban area where voters of the Left parties are used not to be "lured" by theories of "lost" or "useful" vote.

POLLS PRIOR TO ELECTIONS

Opinion polls nowadays tend to cover the observed lack of communication between political parties and social groups. In this way they replace the traditional form of political action through organizations and activists, which were the "sensual nerves" of political parties in the society and were now malfunctioning. This interactive communication is transformed into "expression" of opinion from "representative samples" towards the political parties, which, however, have already taken care of addressing to the civilians, mainly through declarations in the press and the mass media as well in advance. Furthermore, they offer the possibility of examining hypotheses about the political phenomena. It is a fact that 50-60% of the studies published in scientific journals, in North America, Britain and Germany, contains diagrams, equations and tables. The according percentage for the South European countries and those of South America is about 30%, (*c. Riba, European Journal of Political Research Vol. 29 June 1996, pp. 477 – 508*).

Opinion polls, particularly prior to elections, should be used mainly towards two directions. In the first direction that of the percentage of citizens who do not answer, the differences between "intentions of flow" estimated in the tables of flow in opinion polls and the "movement" recorded in exit polls should be examined. At the same time, in pre-electoral polls, depending on the sample used, there is the problem of "answer denial", which explains the underestimation or overestimation from attitude to attitude. The percentage of answer denial, usually reaches 15% –20% and this is not counted in the analysis of the poll. This means that in order to complete 1000 questionnaires about 1200 citizens must be contacted out of whom the 200 do not accept to take part in the research, though the results usually announced concern the 1000 (completed questionnaires). The main problem of the distribution of "undecided" voters, which leads to the inability of referring to the total of the electorate, gives a value to polls concerning the monitoring of attitudes and intentions of the time they are conducted and not a value of "prediction" of the elections result.

In the second direction, the pre-election period between the time of voting and conducting the poll, a time of political activation and intervention of political parties explains any distance between the poll results and the election results. In this respect, analyzing the differences of polls before and after elections can lead to the analysis of influence of the pre-elections period in the formation of the result. It is rather arbitrary to assume that polls conducted 15-20 days before the elections can possibly reach in figures the elections final result, since in this case, what happened between the two measures (before and after) would not have any influence, a result that is arbitrary and wrong at any sense. Pre-election polls should answer central, mainly qualitative, questions. The published pre-election polls did not all record in the same degree the apparent qualitative trend of predominance of PASOK, giving a relevant equality of the two political parties (PASOK & ND). They were rather recording through continuous observations, (mainly during the last ten days before the elections), the final trend of diminishing of (POLAN), though they overestimated DIKKI in 20% of the power of PASOK. Presenting the conducted poll in the A' voting region of Thessaloniki we indicated that:

“Naturally, the course of pre-election period will probably lead to the rearrangement of trends of the electorate. At any rate, the choices of a short pre-election period and the transfer of political activation, almost exclusively in the main political stage, can influence decisively the elections result leading a part of the electorate to different choices other than the two main parties PASOK and ND.”

The above poll (2-6/9/96) was projected in the country and as indicated, during the presentation, it gave an estimation concerning the relation of the two political parties (PASOK & ND) and estimation for the other parties. It is rather common sense to repeat that polls have a value for the political phenomena which they seem to describe, (either as a chronically comparative presentation, or as an observation of their specific features), and that the figures they gave are estimations. In any case, findings of polls should be used having in mind the requirements without trying to exceed the limits of their measurements.