Chapter 4

Modeling Voter preferences

In this study, we want to apply and compare the latest five Greek electoral systems. For this purpose we will use some measures of disproportionality. Each one of these measures uses the election result (seat shares) for the evaluation of a system. For each system, under consideration, one or two election results are available, as each one of the systems, that we study, has been applied once or twice in the Greek Parliamentary Elections. If we want to find a 'good' estimate for a particular measure, and for a particular system, we must use a large number of different election results. Thus, in order to have accurate results, it would be better to use a large number of datasets for each electoral system. These datasets will contain voters' preferences for the alternatives, and thus total votes for each party-candidate.

In this section we deal with the generation of voters' preferences on different alternatives, found in the literature. We examine how we can apply these methods in our study and if it is reasonable enough to use them. Finally, another idea for the generation of Greek electorate will be introduced and will be shown that this is an appropriate method for our study.

Let n be the size of the electorate. Each one of the n voters has to give his preference among m alternatives. We make some assumptions about individual preferences; see, for example, Bordley (1983).

- An individual has preferences among all possible alternatives.
- If an individual prefers alternative 1 to alternative 2 and alternative 2 to alternative 3,

he also prefers 1 to 3. Krantz (1971) reached the conclusion that there exists such a utility function u that whenever the individual finds alternative 1 at least as preferable as alternative 2, we will have $u(\textit{alternative } 1) \geq u(\textit{alternative } 2)$. As the election offers the voter m alternatives, in the election, he will vote in such a way as to maximize his overall utility.

- Sincere-Honest voting: In any electoral system, many voters try to take into account the expected candidate strength, as well as their own preferences, in determining how to vote. Thus, their vote depends on their assessment of how other citizens intend to vote. Here we assume that the voters have no knowledge of the likelihood of success of the various candidates. Then the voter is said to be making a decision under uncertainty. Under uncertainty, the optimal strategy involves voting for one's favorite; see Merrill (1985). In simple words we assume that the voter cast his vote for the alternative to whom his preference is higher.

Given that each individual $i, i = 1, 2...n$ has utility $u_i(m)$ for each alternative m there are $n \times m$ values for $u_i(m)$. Those values consist the *utility matrix* U.

$$U = \begin{bmatrix} u_1(1) & ... & u_1(i) & ... & u_1(m) \\ u_2(1) & ... & u_2(i) & ... & u_2(m) \\ ... & ... & ... & ... & ... \\ u_j(1) & ... & u_j(i) & ... & u_j(m) \\ ... & ... & ... & ... & ... \\ u_n(1) & ... & u_n(i) & ... & u_n(m) \end{bmatrix}$$

The i^{th} row represents the i^{th} voter utilities for the m alternatives. The j^{th} column represents the utilities of all voters for the j^{th} candidate. If U is known, voter preferences among all possible alternatives will be known. Thus, election space will be known.[1] The utility matrix allows us

[1] Camperlin and Cohen (1978) use the term *election* to define the relative frequencies in an electorate of support for each possible permutation (rank ordering) $= m!$ of the m alternatives. The *election space* is the set of all possible elections that could occur when there are m alternatives. It correspond to all points in the space of $m!$ dimensions that have no negative coordinates and have coordinates summing to unity: $m!$ *dimensional simplex.*

Suppose a voter with preference ordering ABC. For each one of the alternatives A, B and C has a preference to know each individual vote, in this particular election with m alternatives. However, the alternatives that appear in a given election are not known before the election and thus utilities are random events. It is equivalent to make the elements of U random variables. But if the utilities are random variables we have to specify their distribution. In other words we need the probability density function on the election space.

4.1 Impartial culture-Uniform assumption

The most prominent efforts for the definition of the probability density function on the election space comes from the impartial culture assumption. The individual $i\prime s$ utility for alternative m can be thought of ranging from strong preference to strong aversion:

$Dislikes\ m$			$Likes\ m$
$Strong$ $Aversion$	$Mild$ $Aversion$	$Mild$ $Preference$	$Strong$ $Preference$

Suppose that the distribution is symmetric: The probability that an individual has a mild preference for m equals the probability that he has a mild aversion for m and the probability this individual has a strong preference for m equals the probability that he has a strong aversion for m.

Moreover, the uniform assumption demands that the chance an individual feels strongly about m equals the chance that he feels mildly about the same alternative. Under this assumption it is reasonable to model the utilities as *uniform random variables.* Thus, this method treats each voter as a sample point drawn from an infinite population in which all possible permutations are equally likely and it is equivalent to assume that all alternatives are equally attractive. For example, the probability of an individual i having various utilities for m is given

($utility$) such that $u(alternative\ A) > u(alternative\ B) > u(alternative\ C)$. For the voter with preference ordering BCA it is $u(alternative\ B) > u(alternative\ C) > u(alternative\ A)$,etc in Figure 1.

Merrill (1984,1988) chose the utilities to be independent, uniformly distributed random variates on a common interval from 0 to 1, in order to generate random society.

We point out that if we generate an electorate with this method and then compute the percentages of votes $v_1, v_2..., v_m$, for each alternative, the above method gives that $P(v_1) = P(v_2) = ... = P(v_m)$, which does not exist in reality. Although it is an important drawback, at first thought it leads us in the decision that this method must be taken into account, as it is the first and the most fundamental method for the generation of voter preferences. Furthermore, Merrill (1984) used the impartial culture and the spatial model in order to compare some well known electoral systems, and he obtained almost the same results.

As, our purpose is to study the results taken from different electoral formulas and not to study voter preferences and to model voter's behavior, a first thought was that it would be reasonable to use this method.

Ludwin (1976) compared six electoral systems (Plurality, Runoff, Alternate, Exhaustive, Alternate, Borda) based on Condorcet winner. In order to do this comparison, he modeled and he simulated a three candidate single-seat election many times. Ludwin(1976), instead of simulating voter utilities, simulated the frequencies of the possible permutations of the three candidates. As three candidates take part in his hypothetical election there are 3!=6 different permutations $(abc, acb.bac, bca, cab, cba)$. Six uniformly distributed pseudo random numbers generated the proportion of the electorate, which is matched with a permutation in the following way. The six random numbers were summed and each one of them was divided by the sum of all. Let $x_1, x_2, ..., x_n$ be the generated pseudo numbers. The $x_i \ / \sum\limits_{i=1}^{6} x_i,\ i = 1, 2, ...6$ consist of the six percentages of permutation. As the percentages of each permutation comes from the uniform distribution, all possible permutations are equally likely . It is equivalent to the assumption that all candidates are equally attractive from voters. Consequently, it is an other way of making the uniform assumption.

4.2 Application of the impartial culture

First of all, we have to point out that Merrill (1984, 1988) as well as Bordley (1983) used this method for the comparison of electoral systems such as Approval voting, Borda, etc.c, where voters have to decide between single candidates and not between both political parties and independent candidates. Furthermore, these systems do not demand the division of the electorate in districts and finally the result of the procedure is the number of votes obtained by each candidate and not the number of seats in the parliament obtained by each party (or independent candidate, if he receives a seat).

As we mentioned before, the main drawback of this method is the assumption that all alternatives are equally attractive. Thus, the generation of voter's preferences for an electorate based on the uniform distribution, leads to a generation of a number of votes obtained by each alternative that is almost the same. In the case of a small electorate, it is reasonable that the variation in these numbers would be sufficiently large. However, in a case of a large electorate, those numbers would be almost the same.

Greek electoral systems demand the generation of a large electorate, as we have to generate voter's preferences for each lower district separately, and the number of them is large (56). After all, the 300 seats correspond to an electorate that ranges from 6 up to more than 8 million people, and an electorate of such range must be used if we want to use the real number of seats. The application of this method gave 56 districts where in each one of them all political parties obtained the same percentage of votes. Table 4 presents the results of the above application, in the case of a small electorate $(n = 1000)$. Table 5 presents the results of the above application in the case of a large electorate $(n = 6000000)$ indicating that the uniform assumption is unrealistic and not suitable for our study.

Table 4: The results for the first 8 lower districts in case of 1000 voters and 5 political parties for the uniform distribution $U(0,1)$. m: political parties, l: lower districts.

$m \backslash l$	1	2	3	4	5	6	7	8
1	17	14	8	4	2	3	3	3
2	13	39	6	2	9	1	2	2
3	11	24	4	5	2	1	4	6
4	14	33	8	5	4	5	9	3
5	8	15	6	8	10	2	3	4

Table 5: The results for the first 8 lower districts in case of 6000000 voters and 5 political parties for the uniform distribution $U(0,1)$.

$m \backslash l$	1	2	3	4	5	6	7	8
1	76211	150516	39141	29152	32406	14818	25638	21821
2	75800	150093	38899	29476	32619	14881	25593	22020
3	75554	150122	38815	29378	32670	14959	25600	21788
4	76025	150512	38909	29570	32341	14798	25432	21594
5	75610	149957	39236	29421	32564	14944	25537	21799

When specific candidates compete it is likely for them to obtain equal number of votes.

4.3 Normal assumption

The assumption that the distribution is symmetric remains. That is, the probability that an individual has a mild preference for m, equals the probability that the individual has a mild aversion for m and the probability that the individual has a strong preference or a strong aversion are identical. In other words, P(mild preference) = P(mild aversion) and P(strong preference) = P(strong aversion). Furthermore, it is assumed that the chance an individual feels strongly about m equals the chance that he feels mildly about him. Under this assumption it is reasonable to model the utilities as normal random variables. The probability an individual i having various utilities for m is given in Figure 2.

It seems that with this method we surpass the main problem of the uniform assumption, all alternatives are not assumed to be equally attractive. This method was used by Bordley (1983) for the comparison of six well known systems. Those are: Borda, Copeland, Approval, One person/One vote, Dictatorial and Random choice system, systems where single candidates compete.

4.3.1 Application of the Normal assumption

We applied this method by using a Normal distribution with zero mean and standard deviation equal to 1, $N\ (0,1)$. Voters' preferences were generated by this distribution. By using this assumption the percentages of all political parties were almost the same in each district, when the size of the electorate in each district was very large. Table 6 presents the results of this simulation. On the other hand, we notice a small difference in the percentage of total votes of the political parties in each district when the size of the electorate was small. Table 7 shows these results. As we have already explained in order to allocate 300 seats in 56 lower districts, to a number of political parties, we have to use an electorate of a large size.

From the Table 6 and the Table 7 is obvious that the Normal distribution is not a suitable distribution for the simulation of the Greek electorate, because the chance that all political parties are almost equally attractive, is unrealistic.

Bordley (1983) used both methods in his study (Uniform and Normal assumption) and he found out that the result does not change a lot as we move from communities with uniform distribution to communities with normal distribution. Bordley (1983) after this reached the conclusion that the utility distribution may not matter too much.

Table 6: The results for the first 8 lower districts in case of 6000000 voters and 5 political parties for the normal distribution $N(0, 100)$.

$m \backslash l$	1	2	3	4	5	6	7	8
1	75694	150212	38927	29166	32710	14859	25643	21785
2	76008	150299	38959	29401	32412	15030	25365	21798
3	75984	149811	39055	29595	32604	14982	25446	21790
4	75807	150306	39010	29381	32465	14742	25515	21956
5	75707	150572	39049	29457	32409	14787	25831	21871

Table 7: The results for the first 8 lower districts in case of 1000 voters and 5 political parties for the normal distribution $N(0, 100)$. m: political parties, l: lower districts.

$m \backslash l$	1	2	3	4	5	6	7	8
1	1	10	2	4	0	1	2	1
2	4	6	3	1	2	1	1	3
3	4	7	1	0	4	0	0	1
4	9	5	2	0	2	1	1	0
5	0	9	1	2	0	2	2	0

4.4 Spatial model

Originally, it was developed in the field of economics by Hotteling (1929) who was interested in explaining market locations. The essential model was developed, independently by Coombs (1950) a phycologist interested in individual choice behavior. It was first used to represent electoral competitions by Downs (1957) and have been studied extensively in the past decade.

The main ideas of the Spatial model are:

- Each voter can be represented by a point in the hypothetical space, so that the point reflects the person's ideal set of policies.
- The policy position can be represented by a point in the same space.
- A voter chooses the candidate whose policy position is closer to his position in the hypothetical space.

Models of this type are also called *Proximity models,* because preference follows directly from 'closeness'. From the above is obvious that the spatial model assumes that both the voters and the alternatives to be distributed in the same space, either unidimensional or multidimensional. Each dimension represents a specific issue. Such issues might be sectors such as political ideology, etc. The voter's and the candidate's positions in the space, depends on their own perception on each one of the issues.

In order to present the main assumptions of the model we give the following multidimensional notation: $x = (x_1, x_2, ..., x_n)$ is the representation of a voter preferences for all issues. Thus, x_i is the voter's position on issue i (dimension i), $i = 1, 2, ..., n$, where n is the number of issues taking into account, $v_j = (v_{j1}, v_{j2}, ..., v_{jn})$ is the position which the candidate j advocate, $j = 1, 2, ..., m$, where m is the number of the candidates. This vector is called candidate j's strategy.

Assumption 1: Although a complete description of the electorate requires several dimensions, a voter may be interested in only some of these dimensions. Hence, he does not assign values for all elements of x. We assume, nevertheless, that every element of x has a numeric value. Every voter acts as if he gives his preferred position on each issue.

Assumption 2: We assume that the vectors x, v_j are continuous variables, although there are issues where preference on them implies continuous or discrete values. The continuous assumption is convenient because it facilitates the use of continuous calculus.

Assumption 3: We assume that all voters act as if they make identical estimates of v_j, which implies that the voters' perceptions are the same for each candidate position.[2] This assumption implies that citizens estimate a position for each candidate on every dimension.

[2] Analysis of real data indicated that these is a substantial disagreement between different individual perception's of candidates. Aldrich and McKelvey (1977) gave an estimation of the true candidate positions, by the positions each voter reports, for each candidate.

Assumption 4: Voter preferences are characterized by a density function $f(x)$. Candidate strategies are represented by points in the coordinate space of this density. The number of densities that might characterize the preferences is infinite. It is assumed that $f(x)$ is symmetric and unimodal.

This model makes some other, more complicated, assumptions about citizen's evaluation of candidates' strategies.For details see Hinich and Ordeshook (1970).

In order to clarify the above model we give a simple example. Voter i has to rank issues in a pre-specified interval scale . If we work in the two dimensional space and gives for issue 1 the value x_i and for issue 2 the value y_i , his position in the two dimensional space is given by the point with coordinates $(x_i,\, y_i)$. Suppose that alternative A is described by the point with coordinates $(x_A,\, y_A)$ and B is described by the point $(x_B,\, y_B)$. According to this model the voters chose the candidate whose issue position on the space is the closest to his issue position. The most common utility function for measuring the 'closeness' is the Euclidean distance d. Thus, voter's i distance from alternative A is given by $d_A = \left[(x_i - x_A)^2 + (y_i - y_A)^2\,\right]^{1/2}$ and his distance from B is given by $d_B = \left[(x_i - x_B)^2 + (y_i - y_B)^2\,\right]^{1/2}$. Furthermore the model assumes that the voter's utility for the candidates decreases linearly with the Euclidean distance d. Thus, if $d_A \geq d_B$ then i prefers B to A. Thus, the use of this special utility function (Loss function) assigns to the voter a single comparable measurement of the phycological distance between his point of view and the location of each alternative. In this model as well as in the uniform and in the normal assumption utilities are used only to determine preference order.

Champerlin and Cohen (1978) used this classical approach to issue voting and generated voter and candidate positions in the space via simulation. Voters were represented in the four dimensional space. Thus, for each voter four numbers had to be simulated representing voter's position in the four dimensions-issues. The four numbers for voter i were generated in the following way:

- Voter i's position on the first dimension Y_{1i} was chosen from the standard Gaussian distribution with zero mean and standard deviation equal to 1.0
- Voter i's position on the second dimension Y_{2i} was generated from the position in the first dimension by perturbing it with Gaussian noise of zero mean and standard deviation

equal to 1.4

- The third was produced from the second with fresh noise of similar character
- The fourth coordinate was generated from the third in the same fashion.
- The values on all dimensions were normalized in order to have unit variance

Also Merrill (1984), (1985) used the spatial model in his study so that both voters and candidates were generated via simulation from the multivariate normal distribution. Although the utility function that he used is the Euclidean distance, he suggested some other Loss functions like

a) *City block metric* given by $\sum |v_i - c_i|$, where $v = (v_1, v_2, ...v_\delta)$ and $v = (c_1, c_2, ...c_\delta)$ are the voter's and the candidate's positions, respectively and δ is the number of dimensions.

b) *Shelpsle utility function* given by $u(d) = exp(-d^2/2)$, where d is the Euclidean distance from voter to candidate. For more details see Merrill (1988) .

c) *Negative of distance* given by $u(d) = -d$, where d is the Euclidean distance from voter to candidate.

Spatial model has been studied a lot and many extensions of it has been found in the literature.. For example, Rabinowitz and MacDonald (1989) gave an extension of the spatial model. Sustaining that the traditional spatial theory of elections is seriously flawed, introduced two new components of issues. Those are the direction and the intensity..

Instead of generating voter preferences for alternatives based on a distribution, we can generate voter preferences for specific issues, using the Spatial model. A hypothetical distribution that estimates the distribution of voter preferences on issues has to be used. This is, usually, the normal distribution, see, Champerlin and Cohen (1978). The use of this distribution leads us to the conclusion that the Spatial model will not produce the 'appropriate' data. These are the data that can be thought as citizen's preferences for political parties and independent candidates, for the Greek electorate.

Consequently, we realize that none of the above methods would be useful in our study.

4.5 The proposed data generation method

Due to the fact that the previous methods are not suitable for our study, we suggest to use real data taken from the latest Greek Parliamentary elections, and to simulate/generate a large number of other dataset by introducing noise (random error) in the initial dataset.

It would be far more interesting to study Greek Electorate in our days, and not its behavior in the past. So a first good step is to use the data from the latest Parliamentary Elections. These are the elections that took place in the year 1996. These data could be easily taken from the Ministry of Internal Affairs. This data set includes vote totals for each political party or independent candidate for each one of the 56 Lower districts. The magnitude for each one of the Lower districts is included. The total number of political parties and independent candidates is 32. In fact the greatest majority of them (26 of 32) obtained a total percentage of votes less than 1%. So it is useless to study each one of them separately. We presume that there are 6 important political parties. The first is P.A.S.O.K with a percentage of total votes in the entire state equal to 41.49%, the second is New Democracy with 38.12%, the third is Politikil Anixi with 2.94%, the fourth is K.K.E. with 5,60%, the fifth is Synaspismos with 5.12, and the sixth is D.I.K.K.I with 4.44%. Despite its small percentage of total votes (2.94%), we include Politikil Anixi in our study, because it is a percentage which allows Politiki Anixi to enter the Parliament only when some of the electoral systems are used. The rest of the parties and the candidates took such a small number of votes that none of the systems under consideration can occupy a seat.

It is very well known that the political party, which governs, for a period of time, losses some voters due to political activities. These voters are divided to other political parties or independent candidates. The second most powerful political party usually increases its total number of votes. The above phenomenon characterizes the Greek electorate. Therefore, it is reasonable to examine the possibility of the number of seats allocated to political parties, by applying the five electoral systems, when we generate datasets according to the characteristic described above. We generate a large number of datasets with the use of the real dataset of total votes, of each political party, in each lower district, and by using the following procedure.

The total number of votes for the first political party, in each district, is eliminated according to a value which is taken from a Normal distribution with zero mean and variance which is related to the real number of votes of this party, in each district. The total number of votes, for the second political party, in each district, is increased according to a value which is taken from a Normal distribution with zero mean and variance which is related to the real number of votes of this party, in each district. For the rest of the political parties we change the total number of votes, in each district, by introducing noise in the initial number of votes. That is, these votes are taken from a Normal distribution with mean value equal to the real number of votes and variance based on the votes, of each political party, in each district. Therefore, the algorithm that we use is the following.

Let n be the real number of districts and m the real number of political parties. For each district i, $i = 1, ..., n$ do

- For the first political party, $m = 1$
  - take a value ε_i from the normal distribution $N(0, c_1\Psi_{1,i})$
  - calculate the absolute value of ε_i
  - the new number of votes for the first political party, $m = 1$, for the i^{th} district is given by $\Psi_{new,1,i} = \Psi_{1,i} - abs(\varepsilon_i)$, where $\Psi_{1,i}$ is the real number of votes of the first political party for the i^{th} district, c is a number which changes the number of votes for the first political party in each district by the percentage we want, and $\Psi_{new,1,i}$ is the generated number of votes for the first political party for the i^{th} district.
- For the second political party, $m = 2$
  - take a value ε_i from the normal distribution $N(0, c_2\Psi_{1,i})$
  - calculate the absolute value of ε_i
  - the new number of votes for the second political party, $m = 2$, for the i^{th} district is given by $\Psi_{new,2,i} = \Psi_{2,i} + abs(\varepsilon_i)$, where $\Psi_{2,i}$ is the real number of votes of the second political party for the i^{th} district, c is a number which changes the number of votes for the second political party in each district by the percentage we want, and $\Psi_{new,2,i}$ is the generated number of votes for the second political party for the i^{th} district.

- For the rest of the political parties, $m = 3, ..., 7$
  - take a value ε_i from the normal distribution $N(\Psi_{m,i}, c_3\Psi_{m,i})$
  - the new number of votes for the rest political parties, $m \neq 1, 2$, for the i^{th} district is given by $\Psi_{new,m,i} = \varepsilon_i$, where $\Psi_{m,i}$ is the real number of votes of the rest political parties for the i^{th} district, c is a number which changes the number of votes for the rest political parties in each district by the percentage we want, and $\Psi_{new,m,i}$ is the generated number of votes for these political parties for the i^{th} district.