Chapter 2

Election Studies and Methodology

Various aspects of the interactions between the electorate and the party system through elections will be analysed in this book. The main points of analysis are as follows: 1) To depict the long term trends in basic voting behaviour and a periodisation of contemporary political history after Independence. The periodisation of contemporary politics has been done by many scholars on the basis of major political events. However, quantitative data can show changing trends more clearly on the basis of mass political behaviour, which is very meaningful in getting a better understanding of the dynamics of Indian politics. And 2) To analyse the relationship between macro socio-economic evolution and electoral participation. The dependent variables concerned are voter turnout and the number of candidates. The former shows the level of participation of the electorate, and the latter the party-people interface. The electoral variables concerning participation seem to be "more basic" than the electoral data concerning party preference. It is more basic because the relations of these variables to socio-economic variables are much closer and more stable over the long term and they are, in that sense, more "embedded" in the society, as will be shown in the analysis. It is important, therefore, to analyse the factors influencing electoral participation especially with reference to the socio-economic variables.

1. Election Studies and Theory of Electoral Participation

In India, many studies have been conducted on every major election. It can be said that all studies on contemporary Indian politics deal in some ways with elections. Studies on domestic politics after Independence, in some

way or another, touch upon the electoral process. In the broadest sense, those studies might be included within the field of election studies.

Election studies started proliferating after the 1960s. The first peak of election studies in the 1960s seemed to have coincided with the dramatic evolution of politics after Nehru's death and the destabilisation of Congress' dominance. The Congress Party's setback in the 1967 general election, both in the centre and major States, appears to have been symptomatic of a transformation in Indian politics. This political situation must have attracted much scholarly attention to the Indian elections, which seems to have been the main factor behind the proliferation of election studies. Since the 1960s, election studies had been produced corresponding to the major elections, especially the Lok Sabha elections. The second peak of election studies seems to have taken place in the 1990s. Scholarly interest in party politics and elections increased after the Lok Sabha election in 1989 because of the collapse of the Congress system and the beginning of the era of the multi-party system and coalition government in the centre. The transformation in the party system seems to be the basic reason why many election studies, especially on the Lok Sabha election, were published.

Studies of the Indian election as a whole do not seem to have kept pace with those in developed countries, especially the United States. In this section, the main studies on Indian elections will be introduced, and theories developed in the United States or in Europe which are relevant in this study will be referred to in order to acquire a better framework for the examination of Indian electoral data.

Election Studies in India

Many election studies in India are based on surveys of individual voters looking into various aspects of the actual voting behaviours, political perceptions and party preferences, socio-economic status, etc. On the other hand, studies focusing on macro patterns or trends in electoral behaviours on the basis of aggregate electoral data, a category into which this study is included, occupy just a minor portion of election studies. Studies based on aggregate election data, and especially those on the Lok Sabha and State Legislative Assemblies, began to come out after the 1970s. Among them, Elkins (1975),[1] Dasgupta and Morris-Jones (1975),[2] Weiner (1977),[3] Blair (1979, 1990 and 1993),[4] Brass (1980 and 1993),[5] Dikshit's edition (1995),[6] Chhibber (1999),[7] and Chhibber and Nooruddin (1999)[8] were major studies.

Elkins' research, carried out on the basis of State Legislative Assembly election data, was early pioneering research work of ecological analysis in

India. Elkins' study based on State Legislative Assembly data from four southern States from 1952 to 1967 showed the importance of the State context, literacy rate, the development of regional communication in explaining turnout and the number of candidates. Dasgupta and Morris-Jones also tried to discover a socio-economic basis for the electoral data. They examined the relations between socio-economic data and election data from State Legislative Assemblies from 1952 to 1967, with the latter including data on electoral participation, competition and party preference. However, their study seems to be unsuccessful in finding a meaningful correlation between the socio-economic data and electoral data, with the exception of participation, namely, turnout.

Dikshit's studies belong to the field of "electoral geography."[9] His studies on Punjab Legislative Assembly elections, using the ecological method provided an interesting analysis of turnout and votes for Congress in relation to the variables of social development, reservation, and linguistic and religious factors. Other interesting ecological studies are those of Chhibber, who insisted that the decline of Congress' electoral fortunes after 1967 was due to the loss of supporters. In another article on the party competition and fragmentation, he showed that the increased competitiveness between parties is not a result of turnout but rather the increased vitality of second parties, and he emphasised the importance of regional context.[10]

There are many studies based on the survey of individual data and case studies of election politics. Those studies in the field include, Weiner and Kothari's edited work (1965),[11] Atal (1971),[12] Indian Council of Social Science Research's edited work (1972),[13] Sirsikar (1973),[14] Kini (1974),[15] Eldersveld and Ahmed (1975),[16] Ganguly and Ganguly (1975),[17] Kaushik (1982),[18] Mayer (1990),[19] Sission and Roy's edited work (1990)[20] including Chhibber and Petrocik (1990),[21] Isaac (1991),[22] Gould and Ganguly's edited work (1993),[23] Meyer (1993),[24] etc.

However, the most important surveys and analysis are those conducted by the Centre for the Studies of Developing Societies (CSDS). The CSDS conducts surveys of individual voters covering all of India. It has conducted its surveys periodically, on the occasion of every major election, since the latter half of the 1960s, with approximately 10,000 samples. It has also published insightful articles and reports based on the surveys.[25] Although the CSDS database seems to be very useful for understanding precise individual behaviours, the individual-level raw data are not available to ordinary researchers.

The scopes and time-periods for the most of these studies, with the exception of the CSDS studies, are, by and large, limited to some States or

areas, and cover a relatively short period. What would be needed in order to understand the basic undercurrent in Indian politics would be a description of the long term and macro trends of voting behaviour among people, and data from the Lok Sabha elections would give ample scope for this. But it is also true that individual-based or case study-based researches, as a whole, constitute a rich depository of information on electoral behaviours. It would be very useful to utilise the information to make up for the shortcomings of studies based on aggregate data, and, in order to check the possibility of ecological fallacy. India is a highly heterogeneous and fragmented sub-continental society. It is, therefore, impossible to ignore the problems of social cleavages, based on caste, religion, class, etc. But usually, it is very difficult for ecological studies based on aggregate data to analyse these problems. It would be productive for the precise information which these individual-based or case study-based researches can provide to be utilised to compensate for the difficulties of estimation on the basis of aggregate data analysis. Such information includes, for example, data on who votes for which party. Individual-based or case study-based researches can effectively check the possible ecological fallacies in this study.

Theory of Electoral Participation

In this section, I will examine studies concerning electoral participation, with the goal of working out a conceptual framework of electoral participation. There are various types of election studies in developed countries. One conspicuous feature seems to be the emphasis on elaborate psychological model-building, based on survey-based data on individuals. A series of theories have been elaborated by scholars, and a theory of multi-dimensional psychological utility space has been developed about how voters choose a candidate or a party. Under this theory, the relative distance and direction between voters and parties in the multi-dimensional psychological utility space are the basic elements in deciding their voting behaviours.[26] However, it can be said that theories concerning electoral participation, or voter turnout, are relatively underdeveloped.

A. Psychological Benefits and Expression

A series of studies by the Survey Research Centre of Michigan University, beginning from the latter half of 1940s, are some of the most important pioneering studies based on surveys of individual voters.[27] The 'Michigan' model emphasises the importance of partisanship or enduring party loyalties in voting behaviour, which they considered the basic element of the

democratic stability of the United States. According to this model, strong partisan feelings contribute to electoral participation.[28] Voters go to the polling booth to express their preferences toward parties and make their favoured parties win. Therefore, the more intense electoral competition between rival parties is, the more people go to vote. This model is still influential among many scholars.

Another important development was the evolution of the spatial theory of party competition, which attempts to uncover factors determining the party preference of voters as well as the conditions leading to electoral participation. This theory began from Anthony Downs' *An Economic Theory of Democracy*.[29] According to Downs, a "rational citizen" makes voting decisions so that the decision brings him/her a maximum psychological utility income under the given conditions. Partisan loyalty is not considered to be a crucial determinant of voting behaviours. V. O. Key also thought that people's partisan loyalties were not unchanging and that people voted, by and large, as subjectively rational citizens on the basis of their responsible appraisals of the policy and performance of the parties concerned.[30] The image of the "rational voter" is modelled as the "utility maximiser model."

Riker and Ordeshook study was the development of the Downs' theory. Their famous formula is as follows:[31]

$$R_i = P_i \bullet B_i - C + D \tag{2-1}$$

where;

- R_i: The utility reward that an individual voter "i" receives from the act of voting.
- B_i: The differential utility benefit that voter "i" receives from the success of his/her more preferred candidate over his less preferred one.
- P_i: The probability that the voter "i" will, by voting, bring about the benefit, B_i.
- C: The cost to the individual of the act of voting.
- D: The basic utility benefits accruing from the satisfaction of voter participation on the basis of the ethics of voting, allegiance to the political system, the confirmation of one's efficacy in the political system, etc.[32]

According to the theory, the probability of the selection of particular candidate and voting for the candidate is higher if the value of "R_i" is higher under the particular set of candidates. In particular, if $R_i < 0$ for any

combination of candidates, then the voter will not participate in the election. The model is a typical "utility maximiser model."

On the other hand, Ferejohn and Fiorina assert that voting decision strategy under high uncertainty does not conform to the utility maximiser model. They evolve a model where voters minimize the probable maximum regret which can accrue from alternative voting decisions. According to this "minimax" regret model, worked out under game theoretic consideration, individuals participate in voting because they are likely to get more regret if their favourite candidate loses the election when they fail to vote for the candidate. The regret is smaller when they vote for the candidate even if the favoured candidate loses the election.[33]

However, these theories on the voting behaviour of "rational" citizens seem to be rather unrealistic. The subjective probability for a voter to have a decisive voice in getting a favourable policy adopted by his/her vote, or an unfavourable policy abolished, is usually extremely low in large constituencies on the basis of the utility maxmiser model. This is because the individual knows that the power of one vote to decide a winning candidate or party is virtually nil. Therefore, "rational" persons do not go to vote because they are essentially unable to influence the result of the election and get benefit from a favourite candidate, unless the positive "D" term is included in the formula (2-1).

The minimax regret model is also unsatisfactory in explaining the psychological process of people's electoral participation. In the case of a contest between two candidates, for example, people are assumed to vote if the cost of voting is less than one-fourth of the amount of the benefit accruing from the victory of their favourite candidate according to their theory. Therefore, if the cost is small enough in a free and fair election, all people will vote, which is a rather unrealistic assumption, especially in huge constituencies. And if the number of candidates is more than two, the calculation of the minimax regret necessitates very complicated game theoretic consideration for voters, which seem usually to be beyond the capacity of ordinary voters.

Thus, these "rational" theories cannot explain fully electoral participation in general. But in one particular election situation, in which voters can feel their votes are "decisive," the utility maximiser model might be more effective. This is the situation in which the expected number of votes of the winner and that of the runner-up are nearly same. In another words, a constituency with an extremely competitive election is likely to see more intense participation by the voters. This is because under such conditions, each voter is in a position to likely decide the winning candidate. The "P_i"

probability, which is a subjective evaluation of whether one will be decisive, will be highest when the number of votes of the first-positioned candidate is expected to be nearly the same as that of the second-positioned candidate. Based on this theory, turnout is expected to be higher when the electoral competition between the first two candidates is intense.[34]

In addition, when intense competition is expected, political parties and candidates will put more effort into mobilising as many voters as possible. In particular, when voting is considered to be a low-cost, low-benefit marginal action for people, the mobilising efforts of strategic politicians, who try to convince people that the electoral race is very competitive, cannot be ignored, as emphasised by Aldrich.[35] These mutually related effects can be called a "party or candidates competition and mobilisation" effect. But the chances of neck and neck competition between first two candidates in election are rather few, even if the competition is subjective evaluation of among people, and, therefore, imaginative one. In the more general situation, the utility maximiser model cannot be realistic unless the "*D*" term is in the formula.

Riker and Ordeshook's "*D*" is, in a sense, considered to be "the expression of the citizen's duty." The psychological utility or satisfaction of citizens can be obtained by the very act of expressing their existence in the legitimate political system through voting, which is not an act based on cost and benefit calculations. People may go to vote in order to express themselves as participants in politics. People get psychological utility by expressing themselves in the form of voting. People might participate in an election even if their favourite candidate has no possibility of winning or even if they know that their vote holds virtually no weight given the huge size of the electorate. People might cast a vote in order to make a statement against the political system, irrespective of who wins or loses. In that case, they might be motivated to vote by the psychological satisfaction that accrues from the expression of their will. They may feel satisfaction from the fulfilment of their duty as citizens, that is, from confirming the legitimacy of the democratic political system through electoral participation. Or they may feel obliged to express their pros or cons opinions on a current political issue or event by voting for a particular party or candidate. These effects are mainly represented in the "*D*" term in the Riker and Ordeshook formula (2-1).

The importance of voting as an expression of the citizen's duty can be clearly demonstrated in several cases from South Asian elections. One notable one is the steady declining trend of voter turnout in Pakistan after the 1980s, as shown in Figure i-1. The decline is considered to be due to the expansion of a general feeling of disappointment among the people toward

the legitimacy of the political system. In another example, the assassination of Prime Minister Indira Gandhi clearly raised the turnout in the 1984 election in India. People went to the polls to express their condolences and sympathy for her son, Rajiv Gandhi. A massive wave of sympathy in favour of the Congress party brought about the wonderful victory for the party.

The theory of voting as a form of "expression" is formulated by Brennan and Lomasky.[36] This "expressive theory" seems to be more useful for backward countries. This is because it is usually much more difficult for the governments in backward countries to provide benefits or rewards to the electorate who vote for the governing parties due to a lack of government capacity. So, a "rational citizen" does not have substantial incentive to go for voting. It would be unrealistic to suppose, in Indian Lok Sabha elections, that people will vote for a particular party with the expectation of any concrete benefit or reward from the party, even when the electoral races are extremely competitive. In any case, according to Brennan and Lomasky, the expressive theory can be formulated as follows:

$$W_i = Y_i + E_i \tag{2-2}$$
$$Y_i = h \bullet (R_{iA} - R_{iB}) \tag{2-3}$$
$$E_i = L_{iA} - L_{iB}, \tag{2-4}$$

where

W_i:	The total expected value for voter "i".
Y_i:	i's expected instrumental value of a vote for A compared to B
E_i:	The expressive value that i places on expressing a preference for A, rather than B.
R_{iA}, R_{iB}:	The benefits that voter i places on A *or* B.
h:	The probability that i will be decisive (i. e., the probability of a tie among all other voters).
L_{iA}, L_{iB}:	The expressive return to i of expressing a preference for A *or* B.[37]

If the cost of voting "C" is combined into (2-2), the condition for participating in an election will be then;

$$W_i > C \tag{2-5}$$

It is obvious that the proportion of Y_i within W_i is usually very small, because h is very small in most cases. It would be especially small in huge constituencies like the Lok Sabha constituencies in India. Anyway, the

substantial portion of "*D*" term in (2-1) is considered to be the same as E_i and h in (2-3) can be supposed to be P_i in (2-1). In addition, E_i can be decomposed into a few components. The following three components cannot be ignored on the basis of the above discussion: 1) The expressive utility difference of the partisan preference for candidate *A* over *B* ($Ep_i = Lp_{Ai} - Lp_{Bi}$), 2) Expressive utility difference between that accruing from the recognition of the legitimacy of the overall democratic as well as electoral system (*C*) and that accruing from the indifference to the political system (*D*) ($El_i = Llc_i - Lld_i$), and 3) Expressive utility difference accruing from the particular issue in the election (*E*) compared to a electoral situation with no conspicuous issue (*F*) ($Ee_i = Le_{Ei} - Le_{Fi}$). Then, the re-formulated equation is:

$$R_i = P_i \cdot B_i + E_i - C + \text{constant} \tag{2-6}$$

Where,

R_i: Personal psychological utility reward

E_i: $= Ep_i + El_i + Ee_i = (Lp_{Ai} - Lp_{Bi}) + (Llc_i - Lld_i) + (Le_{Ei} - Le_{Fi})$ (2-7)

This is a simple formula of the psychological mechanism of electoral participation based on the theories of Riker and Ordeshook, on the one hand, and Brennan and Lomasky on the other. The bigger R_i, the higher the probability that a person "*i*" will cast his/her vote for that particular party. It must be noted that R_i, B_i, and E_i are all difference terms. R_i is the difference of the psychological utility reward of the voter *i* between two situations.

B. Political and Socio-Economic Environment

The realisation of voting through the mechanism R_i is not isolated from the situation surrounding the voter. R_i is a model of internal psychological process. R_i itself does not motivate or facilitate the person to go to the polling booth. We must add more relevant variables in order to get a more realistic conceptional framework. Although R_i is a psychological mechanism, it requires a motivation and environment to work. The most relevant inputs from the outside toward R_i might include "political motivation (*PM*)," "socio-economic personal situation (SP_i)," and "socio-economic environment (*SE*)." The "Socio-economic personal situation (SP_i)" is a personal situation for the person, and differs from person to person. On the other hand, the "socio-economic environment (*SE*)" is a macro socio-economic structure surrounding persons, and is a common environment for the majority of the electorate.

First, political motivation can be defined as a motivational factor causing

politicisation, leading to electoral participation. It includes various aspects of politics, both long-term and short-term, such as the basic confidence in the political system, the situation concerning competition and mobilisation by parties or candidates, current political issues, and political economic issues such as recession, inflation, and unemployment.[38] Electoral participation cannot take place without political motivation. For example, people will not participate in an election if they do not have a minimum basic confidence in the political system. Or a particular ideology such as communism or nationalism may encourage those believing in the ideology to vote for a specific party. Alternately, a particular current issue, such as accelerating inflation or communal riots, will raise the possibility of people voting in order to express their opinions. In short, electoral participation cannot happen without these political motivational inputs.

Secondly, the personal socio-economic situation may be important. The theory that social stratum, economic class, community and other demographic characteristics are important determinants of political behaviours, including voting, has been traditionally emphasised by various scholars in political science in the developed democracies, including Seymour Martin Lipset, Samuel Huntington, G. Bingham Powell, etc.[39] The personal socio-economic situation is not a motivational factor. But it can be an important factor conditioning or facilitating voter participation. For example, illiterate or poor people might be less sensitive to political motivation (*PM*), because their capacities for political communication and perception might be at a lower level, and hence their turnout lower. The poor might be less likely to participate in state institutions such as the electoral system because they are not beneficiaries of government policies.

Contrarily, the personal socio-economic situation of an individual might not be important for electoral participation. Illiterate or poor people might also participate as much as literate or rich people. It is matter of examining the data. According to Indian election studies, based on surveys of individual voting behaviours, there seems to be no clear and consistent evidence that the level of education or economic status of an individual voter are important factors determining electoral participation. Rather, the roles of these variables are ambiguous according to studies based on surveys of individuals.[40] This is very important point, and deserves further consideration.

Table II-1 shows the sense of efficacy of voting[41] among 3,800 and 9,614 respondents surveyed in the 1971 and 1996 Lok Sabha elections by CSDS. The sense of efficacy of voting is an important psychological basis for electoral participation. This sense was generally high and stable, and rose

Table II-1

Efficacy of Vote

	1971	1996
Has effect	48.5	58.6
Makes no difference	16.2	21.3
Don't know	35.3	19.1

Vote has effect	**(%)**
Illiterate	47.0
Scheduled Tribe	47.8
Very poor	50.4
Female	50.8
Aged 56 years or above	51.9
Rural	56.9
Other Backward Class	58.0
Hindu	58.0
All India average	**58.6**
Scheduled Caste	60.0
Muslim	60.3
Aged 25 years or less	60.8
Upper Caste	61.5
Upper class	62.1
Urban	64.1
Male	66.2
Christian	66.4
College and above	79.6

Source: Singh, V. B. 1997. *Elections and Social Change in India: Results of National Election Study, 1996*. New Delhi: Centre for the Study of Developing Societies, p. 65

Note: N=3800 in the 1971 survey and N=9614 in the 1996 survey. Respondents were sampled nationally on the basis of a multi-stage stratified random sampling procedure.

from 1971 to 1996. The stable and extensive existence of the sense of efficacy for voting is an aspect of their confidence toward political system in general and the institutional set-up of election in particular. The lower part of the table shows that the socio-economic background of respondents makes a difference in the degree of sense of efficacy. Socially and educationally higher strata of people feel greater efficacy than the lower ones such as the illiterate, the STs, the poor, female, the old, rural folk, etc. There are other

micro-level studies, based on surveys of individuals, indicating approximately the same tendency.[42]

However, the feeling of efficacy by the individual does not automatically lead to voting. Table II-2 shows the distribution of voters and non-voters by respondent background, on the basis of the same CSDS survey. It is important to note that there is no fundamental correspondence between the respondents' sense of efficacy and the act of voting. Illiterate persons, for example, have almost the same tendency to vote as the all India average. Those with educational attainment of college and above have the lowest level of voting. On the other hand, studies based on aggregate electoral data, both in the developed countries as well as India, show a fairly high correlation between turnout and development indicators. These seemingly contradictory tendencies can be conciliated if we assume that the overall socio-economic environment is essentially more important than the individual socio-economic situation when looking at voter turnout. In other words, it might be that the macro social relations or context are far more important than personal attributes or the socio-economic position of each person, which is the next point to be mentioned. In any case, individual-based data on the individual socio-economic situation cannot be dealt with in this study.

Table II-2

Distribution of Voters and Non-voters by Respondents

Background	Voter (%)	Non-voter(%)
College and above	82.5	17.5
Urban	84.0	16.0
Female	85.1	14.9
Young less than 25 years	85.6	14.4
Upper Caste	85.6	14.4
Aged more than 56 years	85.8	14.2
Muslim	86.3	13.7
Scheduled Tribe	86.5	13.5
Hindu	87.3	12.7
All India Average	**87.3**	**12.7**
Illiterate	87.6	12.4
Other Backward Class	88.2	11.8
Rural	88.3	11.7
Scheduled Caste	89.2	10.8
Male	89.5	10.5

Source: Same as Table II-1, p. 25
Note: Same as Table II-1.

Thirdly, the variables of the socio-economic environment (*SE*) are basically macro indicators for the local society as a whole and are, by and large, common to all persons belonging to the society. They are indicators of the features of the social relations or context. Highly literate societies or areas, and better economic societies or areas, are likely to provide a better environment for individuals to participate in modern governmental institutions, including elections.[43] A minimum level of educational and economic conditions for the society or area is, in a sense, the basic socio-economic infrastructure facilitating electoral participation. It can enhance the capacities of the people to participate in elections. The socio-economic environment includes other aspects such as urbanisation, macro social cleavages, etc. These aspects are not necessarily related to the concept of "development," but may be important in defining the macro social context for people. It is anticipated that these aspects can condition the voting behaviours of people in some way or other.

In this connection, the role of the community or group may also be important. For example, those belonging to a community alienated from the majority community might show consistently higher or lower turnout according to the status of the minorities in the society. The consistency of voting behaviour among members of a community may depend upon the cohesiveness of the community. Theoretically, although the vote of an atomised person is virtually meaningless and he or she will not vote in order to get practical benefits, according to the utility maximiser theory, the batch of ballot papers of a cohesive community or group as a whole is very significant for every candidate if the leader of the cohesive collectivity can deliver the ballot papers to the candidate.[44] A member of the collectivity can know or expect whom the majority of the other members will vote for under the influence of the leader. Their expectations can be amplified under such leadership, and hence, the expected personal utility might be amplified accordingly. This "amplifier effect of the cohesive group" may be considerable in a society where there are severe social cleavages, like India. Some macro demographic indicators might be important.

Finally, one point might be added to clarify the feature of voter turnout. According to the CSDS survey, compared to the stability and extensiveness of the sense of efficacy of voting, the sense of usefulness of political parties among people[45] was at a low level from 1971 to 1996. Table II-3 shows the sense of usefulness of political parties. People believe in the electoral system much more than in the usefulness of the parties. If they do not give high appraisal to the usefulness of political parties, the strength of their partisan preferences as a whole is also expected to be rather weak. It follows that the

Table II-3

Usefulness of Political Parties

Response	1971	1996
Good deal	10.9	9.5
Somewhat	21.6	33.0
Not much	25.7	27.2
Don't know	41.7	30.3

Source: Same as Table II-1, p. 67
Note: Same as Table II-1.

aggregate votes polled by any party can easily fluctuate from election to election, leading to a much lower correlation between the percentage of votes polled by the party and the socio-economic variables. In this sense, party preference is much less restrained by the macro socio-economic context than the micro personal socio-economic situation. A voter might choose a party irrespective of the macro socio-economic environment. Furthermore, his/her particular socio-economic position in the society might be more important for decision-making in the selection of a particular party. The decision-making process on party selection by voters can be revealed only through the study of cross-sectional individual data.

In any case, there is sufficient reason to insist that voter turnout is more embedded in the socio-economic environment than party preference, on the basis of which higher correlations between turnout and socio-economic variables can be expected.

The model under which the socio-economic environment facilitates, or conditions, electoral participation can be called the "socio-economic environment model." Electoral participation under this model is, therefore, considered to be more a socio-economically embedded human behaviour than the rational economistic "calculations" or psychological satisfaction of individual voter. Thus, voting behaviour is closely related to macro socio-economic development, demographic features, social cleavage and hierarchy, etc. The most important focus of this study is the analysis of this aspect.

C. The Conceptual Model

Up until now, we have discussed that both the "Political motivation" and "Socio-economic environment" are essential elements for voting, which is realised through a personal psychological utility reward mechanism (R_i). In the absence of the political motivation, voters do not have the will to go to

the polls. Without a socio-economic environment facilitating or conditioning voting behaviour, voting cannot take a concrete form, even if there is a political motivation on the part of the voter. Moreover, people do not vote without a personal psychological utility reward even though there is political motivation and a socio-economic environment conducive to electoral participation. On the basis of this discussion, the probability of voting (V_{ki}) by a person i in k district might be modelled in a relatively simple way as follows;

$$V_{ki} = f(R_{ki} \bullet PM_k \bullet SE_k) \tag{2-8}$$

Where, $f(x)$ is a monotonically increasing function of variables x.

Here R_{ki}, PM_k, and SE_k, can be decomposed into the average for the country as a whole and the deviation from it. That is;

$$R_{ki} = \overline{R} + \Delta R_{ki},\ PM_k = \overline{PM} + \Delta PM_k, \text{ and } SE_k = \overline{SE} + \Delta SE_k$$

Where "Δ" shows the variable as a deviation from the national average. And;

R_{ki}, $\overline{R}$, ΔR_{ki}: Personal psychological utility reward of a person i in k district in the election, the average personal psychological utility reward for all the persons in the nation, and the deviation of a person i in k district from the national average, respectively.

PM_k, $\overline{PM}$, ΔPM_k: The level of political motivation in k district in the election, the average level of political motivation for all the districts in the nation, and the deviation of k district from the national average, respectively.

SE_k, $\overline{SE}$, ΔSE_k: The socio-economic environment in k district in the election, the average socio-economic environment for all the districts in the nation, and the deviation of k district from the national average, respectively.

Then,

$$R_{ki} \bullet PM_k \bullet SE_k = (\overline{R} + \Delta R_{ki}) \bullet (\overline{PM} + \Delta PM_k) \bullet (\overline{SE} + \Delta SE_k)$$

$$= \overline{R} \bullet \overline{PM} \bullet \overline{SE} \bullet \{1 + (\Delta R_{ki}/\overline{R}) + (\Delta PM_k/\overline{PM}) + (\Delta SE_k/\overline{SE})$$

$$+(\Delta R_{ki}/\overline{R})\bullet(\Delta PM_k/\overline{PM})+(\Delta R_{ki}/\overline{R})\bullet(\Delta SE_k/\overline{SE})+(\Delta PM_k/\overline{PM})\bullet(\Delta SE_k/\overline{SE})$$

$$+(\Delta R_{ki}/\overline{R})\bullet(\Delta PM_k/\overline{PM})\bullet(\Delta SE_k/\overline{SE})\}$$

In this equation, $(\Delta R_{ki}/\overline{R})\bullet(\Delta PM_k/\overline{PM})\bullet(\Delta SE_k/\overline{SE})$, are considered to be sufficiently small compared to other variables. Besides, $(\Delta R_{ki}/\overline{R})\bullet(\Delta PM_k/\overline{PM})$, $(\Delta R_{ki}/\overline{R})\bullet(\Delta SE_k/\overline{SE})$, $(\Delta PM_k/\overline{PM})\bullet(\Delta SE_k/\overline{SE})$, can be considered also to be small compared to, 1, $(\Delta R_{ki}/\overline{R})$, $(\Delta PM_k/\overline{PM})$, and $(\Delta SE_k/\overline{SE})$, and, therefore, can be ignored in the normal situation. Then;

$$R_{ki}\bullet PM_k\bullet SE_k \approx \overline{R}\bullet\overline{PM}\bullet\overline{SE}\bullet\{1+(\Delta R_{ki}/\overline{R})+(\Delta PM_k/\overline{PM})+(\Delta SE_k/\overline{SE})\}$$

$$\mathrm{V_{ki}} \approx f[\overline{R}\bullet\overline{PM}\bullet\overline{SE}\bullet\{1+(\Delta R_{ki}/\overline{R})+(\Delta PM_k/\overline{PM})+(\Delta SE_k/\overline{SE})\}]$$

If we can further assume that $f(x)$ is a monotonically increasing linear function of variable x, then;

$$V_{ki} \approx a_1\bullet\Delta R_{ki} + a_2\bullet\Delta PM_k + a_3\bullet\Delta SE_k + \text{constant} + \text{error} \qquad (2\text{-}9)$$

Where,

a_j:$(j = 1,2,3)$ The coefficient reflecting the effect of the variables to the incremental probability of voting, for the country as a whole, which is supposed to consist of the combination of $\overline{R}$, $\overline{PM}$, *and* $\overline{SE}$.

If $\overline{V_k}$ is the average V_{ki} in the k district, that is, if $\overline{V_k} = (V_{k1}+V_{k2}+ \ldots. + V_{knk})/n_k$, where "$n_k$" is the number of electorate in the k district, the expected number of the voters who actually vote is $n_k\bullet\overline{V_k}$ in a binomial distribution. The expected average turnout in the k district, therefore, is $\overline{V_k}$ $(= n_k\bullet\overline{V_k}/n_k)$. Then;

$$\overline{V_k} = a_1\bullet\Delta R_k + a_2\bullet\Delta PM_k + a_3\bullet\Delta SE_k + \text{constant} + \text{error} \qquad (2\text{-}10)$$

Where,

ΔR_k: The average of ΔR_{ki} for the entire electorate $(= n_k$ persons) in the k district.

ΔR_k can be further decomposed based on the discussion leading to (2-6) and (2-7), which are as follows:

$$\Delta R_k \approx \Delta(P_k \bullet B_k) + \Delta E_k, \qquad (2\text{-}11)^{46}$$

And;

$$\Delta E_k = \Delta Ep_k + \Delta El_k + \Delta Ee_k \qquad (2\text{-}12)$$

The cost "C" and the constant terms in the equation (2-6) are eliminated.

Theoretically, a_j consists of the combination of $\overline{R}$, $\overline{PM}$. But these terms cannot be measured quantitatively beforehand. It is because these are basically imagined psychological terms and are very difficult to measure before an election. $\overline{SE}$ is also considered to be the psychological aspect of the socio-economic environment. All we can do is, therefore, to estimate a_j statistically ex post facto. This conceptual model of voting participation is shown in Figure II-1.

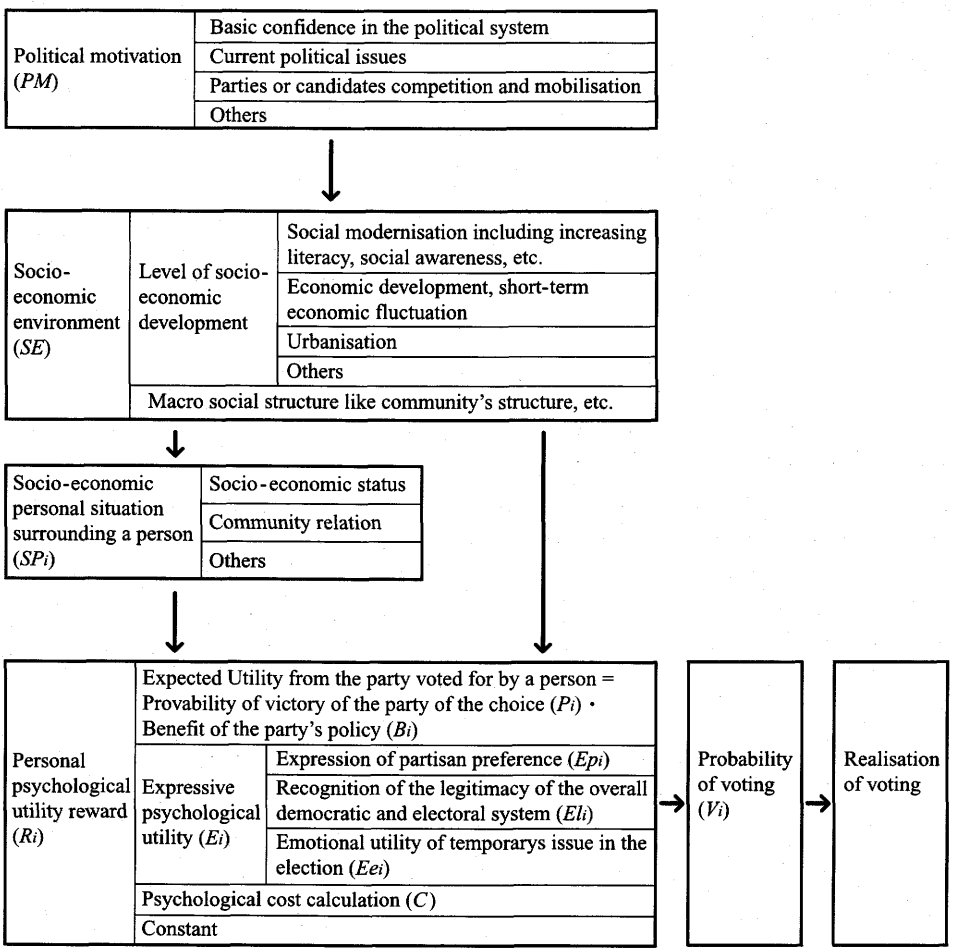
Equation (2-10) is a basic analytical framework for this study. However, it is virtually impossible to estimate "a_1" because we cannot obtain "ΔR_k" data for this study based on aggregate data. The availability of data for "ΔPM_k" is also limited. But the data concerning the "Parties or candidates competition and mobilisation" and "Political situation of each State," in the form of State dummy variables, can be prepared. Besides, in a particular election, the variables representing the "Current political issue" might be able to function as explanatory variables for "ΔPM_k." Finally, for ΔSE_k, I have selected several representative indicators such as literacy and agricultural production, as well as some other demographic data. This could lead to criticism that some of these variables are not appropriate as explanatory variables for electoral participation. My defence is that because these macro socio-economic variables are organically interrelated in the society, it is possible to use them as proxy variables for the "real explanatory variables," which are sometimes imaginary, due to the interrelatedness of the variables.

On the other hand, if, in an extraodinary case, the interaction terms, $(\Delta R_{ki}/\overline{R})\bullet(\Delta PM_k/\overline{PM})$, $(\Delta R_{ki}/\overline{R})\bullet(\Delta SE_k/\overline{SE})$, $(\Delta PM_k/\overline{PM})\bullet(\Delta SE_k/\overline{SE})$, can not be ignored in comparison to, 1, $(\Delta R_{ki}/\overline{R})$, $(\Delta PM_k/\overline{PM})$, and $(\Delta SE_k/\overline{SE})$, the equation (2-1) will be depicted in a more general way;

$$\overline{V}_k = a_1 \bullet \Delta R_k + a_2 \bullet \Delta PM_k + a_3 \bullet \Delta SE_k + b_4 \bullet \Delta R_k \bullet \Delta PM_k + b_5 \bullet \Delta R_k \bullet \Delta SE_k + b_6 \bullet \Delta PM_k \bullet \Delta SE_k + \text{constant} + \text{error} \qquad (2\text{-}13)$$

Figure II-1

Conceptual Model of Voting Participation



Source: Made by the author.

Where;

b_l:(l = 4,5) The coefficient of the interaction term. Theoretically, it includes $\overline{R}$, $\overline{PM}$, and $\overline{SE}$.

This model can be useful for a particular election situation, for example, where a sudden shocking political event, such as the assassination of Indira

Gandhi, happens. In that case, a sudden and overall rise of the level of political motivation can be observed, which means the overall increase of "ΔPM_k". Then "$b_4 \bullet \Delta R_k \bullet \Delta PM_k$" and "$b_6 \bullet \Delta PM_k \bullet \Delta SE_k$" cannot be ignored in (2-13).

Anyway, the equations (2-10) and (2-13), which are illustrated in Figure II-1, are comprehensive framework within which specific or partial models are to be examined in this book.

Concerning the candidates' decisions to run for election, there does not seem to be any firmly established standard theory. This may be partly because party candidates cannot make the decision independently. The party's direction is essential and decisive. We cannot depict the candidate's psychological decision-making process simply as an individual process. In the case of an independent, the candidate has to calculate the "winnability" based on the complex situation of electoral competition vis-à-vis other candidates, popularity among the people, resources like money and supporting organisation, etc. The model becomes extremely complex if it is realistic enough to be able to appropriately anticipate the real process. However, I think the proliferation of candidates in the Lok Sabha elections, especially after 1977, is due to some socio-economic as well as political factors. On the latter point, the situation of the party system surrounding candidates appears to be important. If one party is very influential, other candidates will hesitate to run for election. We can examine these factors in this study.

Concerning party choices of the electorate, in developed countries there are rich accumulations of theoretical studies based on surveys of individuals. However, the main focus of this study is the aspect of electoral participation. Consequently, I do not want to add my attempt to present the theoretical framework of party choice in this study.

2. Ecological Method and Data Processing

In making an analysis of electoral behaviours covering almost all of the main area of India for about a half century, it is inevitable that aggregate data be used. Survey-based individual data are simply unavailable. However, statistical analysis done on the basis of aggregate data sometimes results in a wrong inference unless the data are appropriately processed and statistical methods are applied carefully. In this section, the means of application for analytical methods and data processing will be explained.

Ecological Inference

In this study, an analysis of the long-term trends and patterns of voting behaviour by people in the Lok Sabha elections will be made on the basis of macro statistics. The method using aggregate-level data is called "ecological study."

Ecological studies of election are contrasting with studies of individual-level election survey data. It goes without saying that there are many merits to the analysis of individual-level election data compared to the ecological study. If a questionnaire is well-structured and sampling is conducted properly, the statistical analysis of the individual-level data from the questionnaires can accurately and directly reveal many important factors working at the individual-level. A description of the detailed psychological mechanism of voting behaviours may be constructed on the basis of the analysis, and correlations may be identified between the socio-economic situation of voters on the one hand, and their will to vote or voters' preference for a particular party, and so on. A pioneering research project by the Survey Research Center of Michigan University may be a good example showing the potential of analysing individual-level election data. Their pioneering and classic work made at the end of 1950, *The American Voter*,[47] showed that the partisan loyalties of voters were fundamental and important in explaining their voting behaviours.

On the other hand, there is a serious problem in ecological inference, called the "ecological fallacy." Generally speaking, it is difficult to estimate individual behaviour properly on the basis of aggregate data. This is because detailed information on individuals is lost in the process of aggregation of the data. The problem of ecological inference was clearly pointed out first by William Robinson in 1950.[48]

In India, for example, in some constituencies where the percentage of Muslim population is relatively high, seemingly strange phenomena are sometimes observed where the BJP gets higher percentages of votes in elections when there are communal clashes between Muslim and the majority Hindu community. It is an ecological fallacy if the estimation is made from the observed correlation that Muslim people support BJP more than other people in the communally heightened atmosphere. But in almost all cases, individual level surveys find exactly the opposite, namely that Muslim people usually oppose the BJP more clearly than other religious communities. In the communally tense atmosphere, every community become communally sensitive and tends to support "their" parties. The majority Hindu people are also likely to support the BJP more clearly within

such an atmosphere. That is why the higher percentage of BJP votes is observed. It is due to the fact that demographically, barring Jammu and Kashmir, in most Lok Sabha constituencies Muslims do not constitute a majority vis-à-vis the Hindu community. Even in the Lok Sabha constituencies where the percentage of the Muslim population is much higher than in other constituencies, Muslim people do not constitute a majority. The same thing can happen in interpreting the relation between variables of aggregate voting behaviours and the ratios of SCs or STs in the population. It is difficult to make appropriate inferences even if statistically significant correlations can be found between these variables.

Another major shortcomings of studies based on aggregate data is the limited availability of quantitative data. Usually, it is very difficult to acquire all the relevant data that is needed in the regression. For example, Census reports tell us very little about the political history and social structures that are significant in influencing people's voting behaviours. This is, in a sense, a more serious problem than the ecological fallacy. Although part of the problem can be dealt with on the basis of the "interrelatedness" of macro socio-economic variables, which is discussed just before, the problem of the unavailability of appropriate explanatory variables remains. In order to mitigate the limitation, there is no alternative but to use appropriate "proxy" variables representing the targeted aspects of the social structure.

But as long as it is conducted carefully, ecological inference has some merits vis-à-vis individual-level data. The most important is that aggregate-level analysis will often yield reasonably good estimates of the underlying individual-level effects. The effects of a variable which influences the voting behaviours of all individuals is sometimes more difficult to discern on the basis of the analysis of individual-level cross-sectional survey data than aggregate-level data. For example, macro-economic variables such as income growth, unemployment, and prices, may influence the voting behaviour of all individuals uniformly. In such case, the macro economic variables may not have much impacts in differentiating the individual responses in term of voting behaviours. The macro-economic variable may shift all individuals in terms of voting behaviours, but may not have very different impacts on different individuals.[49] In this case, therefore, it is difficult for individual-level cross-sectional data to provide statistically unbiased and stable estimates of the correlation between an economic impact variable and the voting behaviour.

There are a few ways to reduce the possibility of ecological fallacy while keeping the merits of ecological inference. First, it is important to lower the level of aggregation as far as possible. For example, a village is a better unit

than a block, a block is better that a district, and a district is better than a State as a unit of aggregation in the case of India. In this book, the unit of aggregation is set at the district or, sporadically, Lok Sabha constituency level. It is virtually impossible to collect a complete data set below the district level for all of India from the 1950s to the 1990s. Concerning electoral data, in practice only Lok Sabha constituency level data are available for study. In practice, polling station data or State legislative election constituency data is not available for all of India from the 1950s to the 1990s. Only for the 1999 Lok Sabha election can election data divided on the basis of State legislative constituency be attained from the Election Commission without difficulty. For population census data, of course, detailed decennial data are available down to the village level, but the data cannot be utilised in the regression analysis without matching the variables to other socio-economic or election data. Administratively and historically, the district is a nodal unit based on which the State or central governments are constructed. This increases the importance of the analysis of district based data.

The other way to prevent the ecological fallacy from interfering with the statistical inference or interpretation is to supplement the ecological inference with individual-level survey data. After the 1980s, large-scale opinion surveys have been conducted on a commercial basis. The priority interest of these opinion surveys is which party is winning or losing, the surveys are sometimes very useful, though, because they contain data related to voting behaviours or the socio-economic status of voters. *Market and Research Group—India Today*, for example, has conducted large scale opinion polls at the time of each Lok Sabha election or other nationally important incident since 1984. Although the main focus of the surveys has been to anticipate which parties will win and lose, the survey data also include very useful information on voters, such as socio-economic status, most important interests, views on particular issues at the time of elections, etc. Survey data of this kind can provide very useful supplementary information to prevent the ecological fallacy. On the other hand, there are some limitations to utilising these survey data. Most surveys do not reveal detailed survey procedures such as sampling method, and the results of the survey are shown only in the form of aggregate data. The raw data are usually not available to researchers.

It is rather difficult to confirm the quality of these survey data. Nevertheless, among the surveys covering the whole of India, the series of surveys conducted by CSDS have been highly appreciated. Their data are very useful for understanding individual-level voting behaviours, though the

individual-level raw data are not available to outside researcher. In this book, a summary of the CSDS data will be used to check the ecological inference when necessary.

Variables and Data Processing

This book deals with the 12 Lok Sabha elections from 1957 to 1999 covering major States. The 1952 election is not examined here. It was the only election carried out before the State reorganisation in 1956. That reorganisation brought about major changes of the State boundaries, and they were so extensive that it became very difficult to make comparison between pre-reorganisation and post-reorganisation election data. Though it would be an interesting subject to compare the pre-reorganisation electoral data (=1952 data) and post-reorganisation data in order to evaluate the impact of the State reorganisation, the inconsistency of the data-set makes it extremely difficult.

The data set to be analysed is shown in Table II-4. The data are grouped based on nature and source. Those concerning political motivation (*PM*) include electoral data and State dummy variables. The former data include percentage voter turnout, number of candidates per one million electorate, and percentage of votes polled by major parties such as the Indian National Congress, BJP, etc. It must be noted that in this study, the votes polled by the CPI and CPI(M) are combined after 1964, when CPI(M) separated from the CPI, so that the consistency of the support base has been maintained throughout the period between 1957 to 1999. Dummy variables representing each State are included in order to check whether there are specific factors particular to the State.

Data on the socio-economic environment (*SE*) consist of development related variables and other conditioning variables. The former include percentage of male crude literacy, percentage of female crude literacy, and agricultural developmental data. The latter include demographic data such as percentage of urban population, percentage of cultivators among main workers, percentage of agricultural labourers among main workers, percentage of Non-Hindu population, percentage of SC population, and percentage of ST population.

Among these variables, literacy seems to be a very important indicator of the level of social development. Educational development is one of the most important aspects of social development, but it is not the only aspect of development. The modernisation of society has many facets, including the transformation of feudalistic social structures to a modern society, where

Table II-4
List of Original Variables and Data Sources

		Variables	Name of Variable	Years of Data	Data Source
Political motivation (*PM*)	Electoral data	% of Voter Turnout.	T	57, 62, 67, 71, 77, 80, 84, 89, 91, 96, 98, 99	EC
		Number of Candidates per one million electorate	CAD	57, 62, 67, 71, 77, 80, 84, 89, 91, 96, 98, 99	EC
		% of votes polled by Indian National Congress	INC	57, 62, 67, 71, 77, 80, 84, 89, 91, 96, 98, 99	EC
		% of total votes polled by Communist Party of India (=CPI) and CPI (Marxist)	CP	57, 62, 67, 71, 77, 80, 84, 89, 91, 96, 98, 99	EC
		% of votes polled by Jana Sangh or BJP	BJP	57, 62, 67, 71, 84, 89, 91	EC
		% of votes polled by Janata Party in 1977	JN	77	EC
		% of votes polled by Janata Party in 1980	JJ	80	EC
		% of votes polled by other parties (when necessary)			
	Dummy variables representing States	"AP"= Andhra Pradesh, "AS"= Assam, "BI"= Bihar, "GU"= Gujarat, "HA"= Haryana, "KA"= Karnataka, "KE"= Kerala, "MP"= Madhya Pradesh, "MA"= Maharashtra, "OR"= Orissa, "PU"= Punjab, "RA"= Rajasthan, "TA"= Tamil Nadu, "UP"= Uttar Pradesh, "WB"= West Bengal	AP, AS, BI, GU, HA, KA, KE, MP, MA, OR, PU, RA, TA, UP, WB	57, 62, 67, 71, 77, 80, 84, 89, 91, 96, 98, 99	
Socio-economic environment (*SE*)	Demographic data	% Male Crude Literacy	LM	61, 71, 81, 91, 2001	Census
		% Female Crude Literacy	LF	61, 71, 81, 91, 2001	Census
		% Urban population	U	61, 71, 81, 91, 2001	Census
		% Cultivators / main workers	CW	61, 71, 81, 91, 2001	Census
		% Agricultural Labourers / main workers	AW	61, 71, 81, 91, 2001	Census
		% Non-Hindu population	NH	61, 71, 81, 91	Census
		% Scheduled Castes	SC	61. 71, 81, 91	Census
		% Scheduled Tribes	ST	61, 71, 81, 91	Census
	Agricultural data	Value of 35 crops per hectare (Rs/hectare)	YH	62–65, 70–73, 80–83, 90–93	Bhalla*
		Fertiliser Consumption per net area sown (kg/hectare)	F	62–65, 70–73, 80–83, 90–93	Bhalla*

Sources: "EC" = Election Commission of India; "Census" = Population Census;
* "Bhalla" = Bhalla, G. S. and Gurmail Singh. 2001. *Indian Agriculture—Four Decades of Development*. New Delhi: Sage.

rigid social hierarchies like the caste system become loosened and all kinds of social inequalities are reduced. This kind of social transformation appears to be closely related to changes in gender relations of the society. By and large, social modernisation is clearly connected to gender equality, in which educational equality is the most important aspect. Conversely, educational difference between males and females can be an indicator of the qualitative aspects of social modernisation. If male educational levels far exceed those of females, the society can be estimated to be backward in terms of social structures.

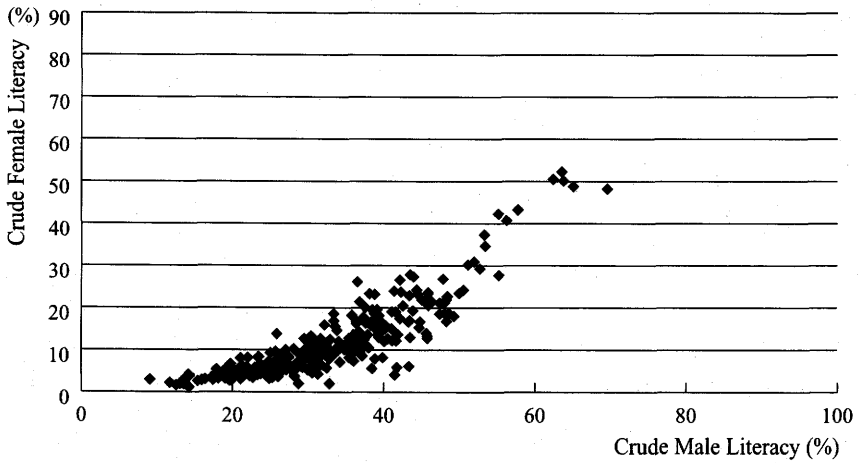
Figure II-2 and II-3 show the relations between male and female literacy in 1961 and 2001, based on district data of the major States.[50] The 1961 graph shows a situation of overall educational backwardness compared to the figure for 2001. It indicates overall backwardness and a skewed social structure. The 2001 graph clearly shows the social development of the following 40 years. Literacy for both males and females has increased, and the gap between the sexes has been reduced. These two variables can be used as proxy variables for estimating the backwardness or skewness of the society, and might reveal interesting aspect of the role of gender in political behaviour.

For the other population census variables, it is rather difficult to anticipate their relation with voting behaviours beforehand. Even when a statistical significance can be found, it may be difficult to properly interpret the explanatory role of the variable because of the possibility of ecological fallacy. However, it is still worthwhile to examine the other variables.

The data on agricultural production are adopted as an indicator of the economic development of districts. This is because agriculture has been and is still the main occupation for many people, with 72.2% still living in rural areas even in the 2001 census. I adopted G. S. Bhalla and Gurmail Singh's compilation of agricultural data, because their data set covers the major States over a period of 30 years for the main crops, in the form of comparable cross-sectional data.[51] From their data set I have adopted two variables: the rupee value of 35 crops per hectare and fertiliser consumption per net area sown. The former variable, which is deflated by the prices in the district, can be an indicator of its economic development, and the latter is a measure of the level of agricultural modernisation. In most cases their data were based on the figures of the central or State governments, but some missing data were estimated. Their data are basically three-year averages for 1962–1965, 1970–73, 1980–83, and 1990–93. These averages were adopted because they tend to increase the confidence limit of the estimates. Short-term fluctuations in agricultural production are offset in the process

Figure II-2

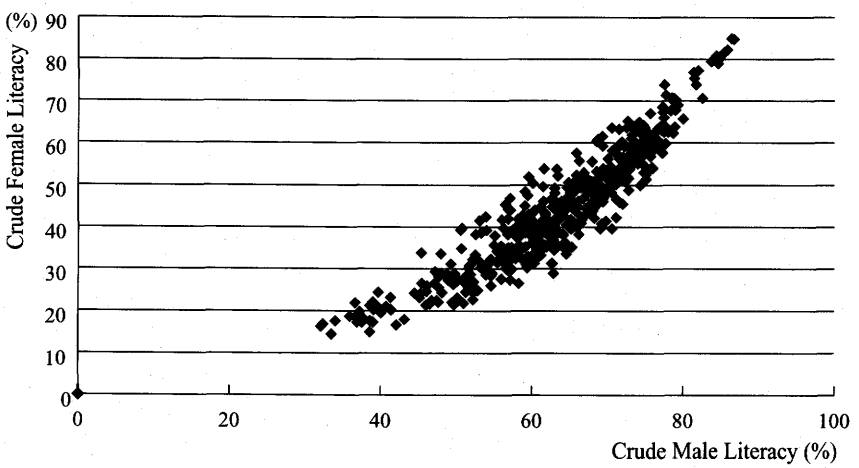
Relation between Male and Female Literacy, 1961



Source: Made by the author based on the data of the 1961 Population Census.
Note: N=303 (districts)

Figure II-3

Relation between Male and Female Literacy, 2001



Source: Made by the author based on the data of the 2001Population Census.
Note: N=492 (districts)

of averaging the figures, and the data set, therefore, is considered to be stable.

The basic unit for Bhalla and Singh's data set is the district. However, the geographical boundaries of districts have changed in many places and on a number of occasions over the 30-year period. In order to maintain a consistency of boundaries for the data set over three decades, they decided to apply the district boundaries from around 1962 to 1965 to the whole data set. They reorganised the agricultural data sets for 1970–73, 1980–83, and 1990–93 to fit the district boundaries of 1962 to 1965, thus creating consistent panel data sets over the three decades. However, there are some areas where the reorganisations of districts boundaries have been so frequent and so extensive, or where there are so many missing values, that data from several districts are combined into a single unit. In central Andhra Pradesh, for example, four districts—Guntur, Prakasam, Nellore, and Kurnool—are combined and taken to be a single unit. The data set does not cover hill or mountain areas, as well as some places where agricultural data are basically not available. In addition, the metropolitan areas of Delhi, Mumbai (Bombay), Kolkata (Calcutta), and Chennai (Madras) are excluded from their data set because their purpose is to analyse agricultural development. However, it covers the major geographical areas of the major States. Consequently, 279 units based on 1962–65 district boundaries are provided. Essentially, the Bhalla and Singh data set is the only one regarding agriculture available for this study.[52]

Urbanisation is another important variable, and is expected to have an important effect on voting behaviour. In classic modernisation theory, urbanisation is assumed to be an important process of social modernisation, which accelerates social mobilisation and changes the pattern of political behaviour. For example, it is considered to lead to a higher level of political participation in various political arenas including elections. In India, like other places, some studies demonstrate that an urban environment brings more exposure to and communication with modern political information, or, more modern political socialisation.[53] But the relative modernity of urban areas has not automatically led to higher turnout. Rather, in a study of turnout in urban areas based on individual survey data, Goel in his study from 1960s data finds that urbanisation did not lead to higher turnout.[54] This is a contradictory finding that goes against classical modernisation theory. We will examine these contradictory arguments in this study.

The most basic problem in applying various data sets for statistical analysis is the adjustment between different sets whose basis of aggregation differ. The geographical boundaries of districts, as mentioned before, have

changed since Independence. In many cases, the figures from the 1961 population census are not comparable with those of 1971 because the boundaries of units, namely, districts, were changed. The same thing can be said when comparing the 1971 census data and that of 1981, and so on. In the case of Lok Sabha electoral data, the constituency boundaries have been reorganised on three different occasions. The first overall delimitation was conducted in 1956, just after the State reorganization. The second was in 1966 and the third in 1976. In 1961, just before 1962 general election, the Two-Member Constituencies Abolition Act 1961 was enacted, and all the two-member constituencies were bifurcated into general constituencies and SC or ST constituencies.

All these inconsistency resulting from difference of geographical boundaries of unit of data must be adjusted for long-term and overall statistical analysis. The adjustment is virtually not possible by aggregating data of smaller unit of each variable along with the fixed boundaries. Lok Sabha election data, for example, can not be compiled along with Bhalla and Singh's district boundaries on the basis of State Legislative Assembly constituencies wise Lok Sabha election data, simply because such detailed data are not available except for 13th Lok Sabha election. In this book a second best method is adopted to get consistent panel data set for all the variables. I have proportionally distributed and aggregated both the election data and the population census data in accordance with the district boundary used in Bhalla and Singh's study shown in Figure II-4 based on the weight of the electoral population of the State Legislative Assembly constituencies included in the district.[55] As explained before, a Lok Sabha constituency consists of the collection of several State Legislative Assembly constituencies. In addition, the Assembly constituencies are usually grouped district-wise according to Delimitation Orders. The boundaries of State Legislative Assembly constituencies, therefore, do not cross the boundaries of districts except a few cases.[56] Figure II-5 next is the illustration of the relationship of these boundaries and the formula of the calculation of the estimated voter turnout in a district. In the figure, there are 1 district and 3 Lok Sabha constituencies consisting of 16 State Legislative Assembly constituencies. For example, the voter turnout adjusted to the district boundary is estimated as follows.

The estimated voter turnout in the district =

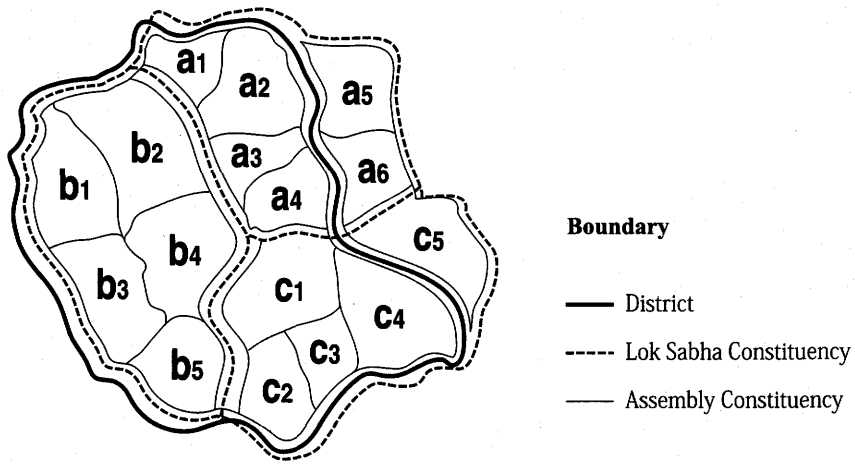
$$\frac{(a_1+a_2+a_3+a_4)\cdot T_a+(b_1+b_2+b_3+b_4+b_5)\cdot T_b+(c_1+c_2+c_3+c_4)\cdot T_c}{(a_1+a_2+a_3+a_4)+(b_1+b_2+b_3+b_4+b_5)+(c_1+c_2+c_3+c_4)}$$

Figure II-4

Geographical Boundary of Units (district or districts combined) Used in this Study

Source: Lines of boundaries are drawn by the author based on the boundary data in G. S. Bhalla and Gurmail Singh, op. cit.

Figure II-5
Adjustment of Election Data



Source: Kondo, Norio. 1997. "Long-Term Trends of Voting Behavior: Parliamentary Elections in India, 1957–91." Journal of the Japanese Association for South Asian Studies (Minami-Azia-Kenkyu [in Japanese]), No. 9, p. 9.

where;

a_i, b_i, c_i = The number of electorate in "i"th Assembly constituency in the a, b, and c Lok Sabha constituencies.

T_a, T_b, T_c = The percentage voter turnout in the a, b, and c Lok Sabha constituencies.

Adjusting the population census data to Bhalla and Singh's district boundaries is not as difficult as it is for the Lok Sabha data. Bhalla and Singh's book contains a table showing combinations of districts which is, by and large, large enough to offset the small changes of district boundaries that took place over the four decades. And in many cases, the changes in district boundaries did not necessitate a proportional distribution as explained above. For example, when a district was bifurcated, what was necessary was to add

up the raw data from the newly formed districts and calculate a percentage ratio. However, for more large-scale reorganisations, a more detailed proportional distribution was adopted to obtain the adjusted value. The Bifurcation of Bombay State into Maharashtra and Gujarat in 1960 and the bifurcation of the old Punjab State into the new Punjab and Haryana are such cases. In these cases the estimations by proportional distribution and aggregation are less accurate. However, in these cases also, the accuracy of the adjusted data does not appear to be bad enough to damage the overall statistical inference in this study, because the district-based aggregate data for the population census is not very different from adjacent districts.

Finally, interpolation or extrapolation is the next step in estimating the census data and Bhalla and Singh's agricultural data in each election year. A simple linear estimation is applied to get the estimated values.[57]

The detailed procedures of the proportional distribution and aggregation are shown in Appendix I.

Notes

[1] Elkins, David L. 1975. *Electoral Participation in a South Indian Context*. Delhi: Vikas.

[2] Dasgupta, Biplab and W. H. Morris-Jones. 1975. *Patterns and Trends in Indian Politics—An Ecological Analysis of Aggregate Data on Society and Elections*. New Delhi: Allied Publishers.

[3] Weiner, M. and J. O. Field. 1977. "India's Urban Constituencies." in Weiner, M. and J. O. Field (eds.). *Electoral Politics in the Indian States* (4 Volumes). New Delhi: Manohar.

[4] Blair, H. W. 1979. *Voting, Caste, Community, Society—Explorations in Aggregate Data Analysis in India and Bangladesh*. Delhi: Young Asia Publication; ———. 1990. "Electoral Support and Party Institutionalisation in Bihar: Congress and the Opposition, 1977–85." in Sission, Richard and Ramashray Roy (eds.). *Diversity and Dominance in Indian Politics, Volume 1, Changing Bases of Congress Support*. Delhi: Sage; ———. 1993. "Local Support Bases and the Ninth General Election in Bihar and Maharashtra." in Gould, Harold A. and Sumit Ganguly (eds.). *India Votes—Alliance Politics and Minority Governments in the Ninth and Tenth General Elections*. Boulder: Westview.

[5] Brass, Paul R. 1980. "The Politicization of the Peasantry in a North Indian State." *Journal of Peasant Studies*, July and September; ———. 1993. "Caste, Class, and Community in the Ninth General Elections for the Lok Sabha in Uttar Pradesh" and "The Rise of BJP and the Future of Party Politics in Uttar Pradesh." in Gould, Harold A. and Sumit Ganguly (eds.). *India Votes—Alliance Politics and Minority Governments in the Ninth and Tenth General Elections*. Boulder: Westview.

[6] Dikshit, R. D. 1995. *Geography of Elections: The Indian Context*. New Delhi: Rawat.

[7] Chhibber, P. K. 1999. *Democracy without Associations: Transformation of the Party System and Social Cleavages in India*. New Delhi: Vistaar, Chapter 5 "Building Coalitions in the Indian States."

[8] Chhibber, P. K. and Irfan Nooruddin. 1999. "Party Competition and Fragmentation in Indian National Elections: 1957–1998." in Roy, Ramashray and Paul Wallace (eds.). *Indian Politics and the 1998 Election: Regionalism, Hindutva and State Politics*. New Delhi: Sage.

[9] Concerning the concept of "Electoral Geography," see, Shelley, Fred M., R. J. Johnston, and Peter J. Taylor. 1990. "Developments in Electoral Geography." in Johnston, R.J., F.M. Shelley, and P.J. Taylor (eds.). *Developments in Electoral Geography*. New York: Routledge.

[10] Chhibber, Pradeep K. and Irfan Nooruddin, *op. cit.*, p. 53.

[11] Weiner, M. and R. Kothari (ed.). 1965. *Indian Voting Behaviour—Studies of the 1962 General Elections*. Calcutta: Firma K. L. M.

[12] Atal, Yogesh. 1971. *Local Communities and National Politics*. Delhi: National Publishing House.

[13] Indian Council of Social Science Research (ed.). 1972. *Studies in The Fourth General Elections*. New Delhi: Allied.

[14] Sirsikar, V. M. 1973. *Sovereigns without Crowns—A Behavioural Analysis of the Indian Electoral Process*. Bombay: Popular Prakashan.

[15] Kini, N. G. S. 1974. *The City Voter in India—A Study of 1967 General Elections in Nagpur*. New Delhi: Abhinav Publications.

[16] Eldersveld, S. J. and Bashiruddin Ahmed. 1978. *Citizens and Politics—Mass Political Behavior in India*. Chicago: University of Chicago.

[17] Ganguly, B. and Mira Ganguly. 1975. *Voting Behaviour in A Developing Society—West Bengal: A Case Study*. New Delhi: Sterling Publishers.

[18] Kaushik, Susheela. 1982. *Elections in India—Its Social Basis*. Calcutta: K.P. Bagchi & Co.

[19] Mayer, Peter. 1990. "The Year the Vote-Banks Failed: The 1967 Elections and the Beginning of the End of Congress Party Dominance." in J. Masselos (ed.). *India: Creating a Modern Nation*. New Delhi: Sterling.

[20] Sission, Richard and Ramashray Roy (eds.). 1990. *Diversity and Dominance in Indian Politics, Volume 1, Changing Bases of Congress Support*. Delhi: Sage; Roy, Ramashray and Richard Sission (eds.). 1990. *Diversity and Dominance in Indian Politics, Volume 2, Division, Deprivation and the Congress*. Delhi: Sage.

[21] Chhibber, P. K. and J. R. Petrocik. 1990. "Social Cleavages, Elections and the Indian Party system." in Sission, Richard and Ramashray Roy (eds.). *Diversity and Dominance in Indian Politics, Volume 1, Changing Bases of Congress Support*. Delhi: Sage.

[22] Isaac, T. M. Thomas and S. Mohana Kumar. 1991. "Kerala Election, 1991: Lessons and Non-Lessons." *Economic and Political Weekly*, 23 November.

[23] Gould, Harold A. and Sumit Ganguly (eds.). 1993. *India Votes—Alliance Politics and Minority Governments in the Ninth and Tenth General Elections*. Boulder: Westview.

[24] Meyer, Ralph C. and David S. Malcolm. 1993. "Voting In India—Effects of Economic Change and New Party Formation." *Asian Survey*, Vol. XXXIII, no. 5, May.

[25] See, for example, Singh, V. B. 1997. *Elections and Social Change in India: Results of National Election Study, 1996*. New Delhi: Centre for the Study of Developing Societies; Articles by the CSDS team in the *Economic and Political Weekly*, 21–28 August and 4 September 1999 on the 12th Lok Sabha election survey conducted by the CSDS; Articles by CSDS team in *Frontline*, 5 November, 19 November, 26 November, and 10 December 1999.

[26] See, for example, Merrill, Samuel III and Bernard Grofman. 1999. *A Unified Theory of Voting—Directional and Proximity Spatial Models*. Cambridge: Cambridge University Press.

[27] Concerning the evaluation of the Michigan model, see, Norris, Pippa (ed.). 1998. *Elections and Voting Behaviour: New Challenges, New Perspectives*. Aldershot: Ashgate/Dartmouth Publishers, pp. xv–xvii.

[28] Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley & Sons, p. 40.

[29] Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

[30] See, Key, V. O. 1966. *The Responsible Electorate: Rationality in Presidential Voting 1936–1960*. Cambridge: Harvard University Press, Chapter 3.

[31] The position of the suffix "*i*" is changed from the original text so that there is no confusion in the discussion below. All the suffixes in the following section are to be attached in the lower-right position.

[32] Riker, William H. and Peter Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review*, Vol. 62, pp. 25–28.

[33] Ferejohn, John A. and Morris P. Fiorina. 1974. "The Paradox of Not Voting: A Decision Theoretic Analysis." *American Political Science Review*, Vol. 68, no. 2, June.

[34] See also, Barzel, Yoram and Eugene Silberberg. 1973. "Is the Act of Voting Rational?" *Public Choice*, Vol. 16; Silberman, Jonathan and Garey Durden. 1975. "The Rational Behavior—Theory of Voter Participation: The Evidence from Congressional Elections." *Public Choice*, Vol. 23; Settle, Russell F. and Buron A. Abrams. 1976. "The Determinants of Voter Participation: A More General Model." *Public Choice*, Vol. 27.

[35] Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science*, Vol. 37, no. 1, February.

[36] See, Brennan, Geoffrey and Loren Lomasky. 1997. *Democracy and decision: The pure theory of electoral preference*. New York: Cambridge University Press.

[37] *Ibid.*, pp. 21–24.

[38] Fluctuations in macro-economic conditions may also affect the turnout. Aguilar and Pacek, for example, insist that "economic downturns lead to increases in electoral participation as those citizens for whom the stakes are highest turn out in greater numbers." Aguilar, Edwin Eloy and Alexander C. Pacek. 2000. "Macroeconomic Conditions, Voter Turnout, and The Working-Class: Economically Disadvantaged Party Vote in Developing Countries." *Comparative Political Studies*, Vol. 33, no. 8, October, p. 1006. But a preliminary examination of their hypothesis on the basis of my panel data set of India's Lok Sabha data does not show any statistical evidence supporting their hypothesis. I, therefore, did not include a examination of their hypothesis in this study.

[39] See, Lipset, Seymour Martin. 1981. *Political Man: The Social Bases of Politics*. Maryland: Johns Hopkins University Press, (First published in 1960). Especially

see, "Chapter 7—Elections: The Expression of the Democratic Class Struggle"; Huntington, Samuel. 1968. *Political Order in Changing Societies*. New Haven: Yale University Press; Powell, G. Bingham, Jr. 1982. *Contemporary Democracies: Participation, Stability, and Violence*. Cambridge: Harvard University Press. Especially see, "Chapter 6—Citizen Involvement / Participation or Turmoil."

[40] See, Varma, S. P. and Iqbal Narain. 1973. *Voting Behaviour in A Changing Society: A Case Study of the Fourth General Election in Rajasthan*. Delhi: National Publishing House, pp. 24–25; Goel, Madan Lal. 1976. "Urban-Rural Correlates of Political Participation in India." in Rosenthal, Donald B. (ed.). *The City in Indian Politics*. Faridabad: Thomson Press, p. 19; Singh, V. B., *Elections and Social Change in India: Results of National Election Study, 1996, op. cit.*, p. 25; *Frontline*. 1999. Heath, Oliver, "Election Analysis: The turnout factor?," 5 November, p. 125.

[41] Aggregation of answers of the respondents for the question, "*Do you think your vote has effect on how things are run in this country, or, do you think your vote makes no difference?*," Singh, V. B., *Elections and Social Change in India: Results of National Election Study, 1996, op. cit.*, p. 64

[42] See, for example, Eldersveld, S. J., and Bashiruddin Ahmed, *op. cit.*, pp. 189–192.

[43] Hadenius' cross-country statistical study shows, for example, such tendency. See, Hadenius, Axel. 1992. *Democracy and Development*. Cambridge: Cambridge University Press. Leighle and Nagler also show the importance of socio-economic status and problems involving its measurement on the basis of data for American presidential elections. See, Leighley, Jan E. and Jonathan Nagler. 1992. "Socioeconomic Class Bias in Turnout, 1964–1988: The Voters Remain the Same." *American Political Science Review*, Vol. 86, no. 3, September.

[44] Concerning the theoretical aspect see, Uhlaner, Carole J. 1989. "Rational Turnout: The Neglected Role of Groups." *American Journal of Political Science*, Vol. 33, no. 2, May.

[45] The aggregation of answers of the respondents to the question, "*How much in your opinion do political parties help to make government pay attention to the people—good deal, somewhat or not much?*" Singh, V. B., *Elections and Social Change in India: Results of National Election Study, 1996, op. cit.*, p. 66.

[46] From (2-6), the average R_i in the k district, that is, R_k is $R_k = P_k \bullet B_k + E_k - C +$ constant. $P_k = \overline{P} + \Delta P_k$, $B_k = \overline{B} + \Delta B_k$, $E_k = \overline{E} + \Delta E_k$ where "–" show the national average. Then;

$P_k \bullet B_k = (\overline{P} + \Delta P_k) \bullet (\overline{B} + \Delta B_k) = \overline{P} \bullet \overline{B} + \Delta P_k \bullet \overline{B} + \overline{P} \bullet \Delta B_k + \Delta P_k \bullet \Delta B_k$. If "$m$" is to be the number of the districts in the country, then, the average of $P_k \bullet B_k$ is;

$$1/m \bullet \sum_{k=1}^{m} P_k \bullet B_k = \overline{P} \bullet \overline{B} + (1/m \bullet \sum_{k=1}^{m} \Delta P_k) \bullet \overline{B} + \overline{P} \bullet (1/m \bullet \sum_{k=1}^{m} \Delta B_k) + 1/m \bullet \sum_{k=1}^{m} (\Delta P_k \bullet \Delta B_k)$$
$$= \overline{P} \bullet \overline{B} + 1/m \bullet \sum_{k=1}^{m} (\Delta P_k \bullet \Delta B_k)$$

Here we may be able to assume that there is no necessary correlation between P_k and B_k in the nation. Then $\sum_{k=1}^{m} (\Delta P_k \bullet \Delta B_k) = 0$, therefore, the average $P_k \bullet B_k = \overline{P} \bullet \overline{B}$

$$\overline{R} = 1/m \bullet \sum_{k=1}^{m} R_k = 1/m \bullet \sum_{k=1}^{m} P_k \bullet B_k + 1/m \bullet \sum_{k=1}^{m} E_k - C + \text{constant} = \overline{P} \bullet \overline{B} + \overline{E} - C + \text{constant}$$

where, $\overline{E} = 1/m \bullet \sum_{k=1}^{m} E_k$

Then;

$$\Delta R_k = R_k - \overline{R} = (P_k \bullet B_k + E_k - C + \text{constant}) - (\overline{P} \bullet \overline{B} + \overline{E} - C + \text{constant})$$
$$= (P_k \bullet B_k - \overline{P} \bullet \overline{B}) + (E_k - \overline{E}) = \Delta(P_k \bullet B_k) + \Delta E_k$$

[47] Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, *op. cit.*

[48] See, King, Gary. 1997. *A Solution to the Ecological Inference Problem—Reconstructing Individual Behavior From Aggregate Data*. Princeton: Princeton University Press, p. 4.

[49] See, Kramer, Gerald H. 1983. "The Ecological Fallacy Revisited: Aggregate- versus Individual-level Findings on Economics and Elections, and Sociotropic Voting." *American Political Science Review*, Vol. 77, no. 1. Kramer recommended the analysis of aggregate-level time-series data rather than individual-level cross-sectional data. However, his argument can apply to the many cases of aggregate-level cross-sectional data vs. individual-level cross-sectional data.

[50] The concept of "Crude literacy" is used in this study. This is in order to minimise the effect of changes of definition of literacy between censuses.

[51] Bhalla, G. S. and Gurmail Singh. 2001. *Indian Agriculture—Four Decades of Development*. New Delhi: Sage.

[52] In July 1995, while querying officials of the Ministry of Agriculture it became apparent that even the Ministry in New Delhi could not provide such a mutually comparable data set for over 20 years. In essence, it was only Bhalla and Singh's study which is available for my study.

[53] See for example, Kini, *op. cit.*; Rosenthal, Donald B. 1970. *The Limited Elite: Politics and Government in Two Indian Cities*. Chicago: The University of Chicago Press; Rosenthal, Donald B. (ed.). 1976. *op. cit.*

[54] Goel, Madan Lal. *op. cit.*

[55] This is the method which R.D. Dikshit recommended for getting an estimation of the value of unit with different boundaries. See, Dikshit, R. D. *op. cit.* Chapter 3 "On the Problem of Unit of Analysis."

[56] Election Commission of India. 1957. *Delimitation of Parliamentary and Assembly Constituencies Order, 1956*. New Delhi; ———. 1967. *Delimitation of Parliamentary and Assembly Constituencies Order, 1966*. New Delhi; ———. 1976. *Delimitation of Parliamentary and Assembly Constituencies Order, 1976*. New Delhi: Government of India Press.

[57] In the case of Fertiliser, several values in 1957 and 1962 are below zero as a result of the extrapolation by linear estimation. In these cases, the value = 0 is substituted.