National Election Study 2009: A Methodological Note

LOKNITI TEAM

The analysis of the national and state level verdict in the Lok Sabha elections presented in this issue of the *Economic & Political Weekly* is based upon the National Election Study 2009 (nes 2009), a post-poll survey conducted by a team of scholars across the country and coordinated by Lokniti: Programme for Comparative Democracy at the Centre for the Study of Developing Societies (cSDs). Lokniti, while a programme of the cSDs, is also a network of about 40 scholars based at various universities and colleges spread across the country (see www.lokniti.org/index.htm for more details). nes 2009 continues the series begun in 1967 by the cSDs (with a break between 1971 and 1996). The current nes is the largest and the most comprehensive social scientific survey of Lok Sabha Election (see Table 1, p 197), and perhaps of any election in the world. It adheres to the best international practices and protocols of survey research besides evolving new practices for improving the quality and substance of the surveys. This note aims to discuss the salient features of the current nes, its methodological attributes and fieldwork protocols, while placing it in the context of past electoral studies. It will also describe the new features introduced in this round of the survey.

nes 2009 is a post-poll survey, i.e., it is a survey conducted at the place of residence of the respondent after the day of polling. Indian elections allow an unusual and ideal window, in the time period between the end of polling and before the results of the elections are declared, to carry out the post-poll survey. Ever since 1998, this window has been used to conduct the nes. This is distinct from an exit-poll survey in which voters are interviewed outside a polling booth. A post-poll survey aims to provide an insight into voters’ political preferences and their electoral choices rather than to forecast the electoral outcome. While data thus collected can be used for making projections of the likely outcome of the elections, and was indeed used this time, that is not the main aim of this exercise. While the nes gathers robust information about how Indians voted, this is not just a study of voting behaviour. The nes treats elections as a window to capture the most accurate snapshot of the political behaviour, attitudes and opinions of Indian citizens on issues as diverse as the economy, national security, democracy and diversity.

nes 2009 belongs to and builds upon a long tradition. The study of Indian elections based on survey methods began in the cSDs in the 1960s, and led to the first nes in 1967. The nes series, which now includes surveys carried out in 1967, 1971, 1996, 1998, 1999, 2004 and 2009, is the most comprehensive information database of social and political change in India and these surveys have been cited as well as used extensively as a source for political science research (Jaffrelot 1996; Shastri et al 2009; Mitra and Singh 1999; Varshney 2003; Yadav 1999 and 2004; Yadav-Palsikar 2003 and 2009; Special Issues of the *EPW*, 1999 and 2004; *JSPE* 2003; to name a few) as well as research that has covered non-political science related issues (for example, for social mobility studies: Kumar et al 2002a, 2002b; Vaid and Heath 2009 to name a few).

The nes series of surveys can be divided into three distinctive generations. Briefly, the first generation of the surveys was from 1967 to 1971 (with the 1980 pre-poll survey that cSDs designed, but Indian Institute of Public Opinion conducted, falling into this first category). The second generation marked the resurgence of the nes in the 1990s (1996-99 elections) after a gap of two decades. The nes 1996, 1998 and 1999 were conducted on a nationally representative panel of respondents with a total of six waves of surveys (pre- and post-poll). The nes 2009 falls into the third generation of these surveys (which also includes the 2004 nes), which, while employing the best international practices of survey research, has introduced a series of new elements and methodological innovations while also providing a manifold increase of the sample size. Table 1 presents information on the achieved sample size and the number of variables in each of the nes surveys since 1967.

Other than differences in sample sizes and variables in the three generations of the nes, there are other important distinguishing features amongst them as well. The first generation of the nes surveys used a self-weighted national probability sample, where 55 parliamentary constituencies (pcs) were selected by stratifying them on the basis of party competition types. Within these pcs, assembly constituencies (acs) and polling stations (pss) were selected by following the procedure of probability proportionate to size (pps). In the second generation of the nes, again a self-weighted national probability sample was drawn. However, the increase in sample size ensured that this sample was representative for major states, along with the national representation. The pps procedure was followed for selecting the pcs and acs among them. The third generation of the nes surveys, of which nes 2009 is a part, have departed from the previous sampling procedure. Rather than drawing a self-weighted national probability sample, now probability samples were drawn at the state level. These probability samples at the state level were aggregated to provide a representative sample at the national level, thus allowing for detailed...
state-level analysis along with the national analysis of the electorate. All NES surveys (except 2004) have followed the procedure of non-substitution in sampling (discussed below).

In addition to the sampling procedure, many distinguishing features set the NES 2009 apart from the previous election studies. In particular, the use of a “split sample” in which five sets were left out of the sample for logistical reasons. In every other state and union territory, all the PCs were included in the sample.

In the second stage, sampling of ACS within the 536 PCs was done. The number of ACS selected varied from state to state, depending on the overall target sample and the number of ACS within each PC in that state. The sampled ACS within each PC varied from one or two to five or six in medium states and went up to 18 ACS in some of the smallest states. This was done to yield the appropriate number of polling stations and respondents as well as to ensure a good spread of respondents across the entire state. The selection of ACS was done on the basis of systematic random sampling from within all the ACS of that PC. A total of 728 assembly segments were selected, of which we were able to conduct interviews in 722 ACS.

The third stage involved the sampling of PS within each sampled AC. As a rule, four PSs were selected within each AC. The only exception was Uttar Pradesh where the number of ACs was so large that it was not feasible to do more than three PS per AC. The selection of PSs was done by listing all the ps within the sampled AC in the serial order followed by the Election Commission and using the systematic random sampling procedure. As a result 2,808 PSs (typically villages or urban wards) were sampled across the country. This compares favourably with the 2,380 sampling points in NES 2004.

The fourth and final stage in the sampling was the selection of the respondents. This was done by drawing a sample of 25 respondents in urban PS and 20 respondents in rural PS from the latest electoral rolls from each selected polling station. This increase in target sample to 20-25 in each PS, when compared to the 15 from each PS in the 2004, was made to ensure not only a larger achieved sample, but also a better spread of the data. The number for rural and urban areas was decided on the basis of differential achievement rates recorded in the past. Electoral rolls of the sampled PSs were obtained from the website/office of the Chief Electoral Officer of the state. In every polling station the requisite numbers of names were selected from the electoral rolls by systematic random sampling procedure.

In each sampled polling station area, field investigators were given a list of sampled respondents containing the name, age, gender and address of each of the sampled respondents and were asked to approach them. Thus 59,650 names were asked as the targeted sample for NES 2009.

The investigators were asked to interview only those whose names were provided to them. After a review of the experience of partial substitution in NES 2004, it was decided that substitution was not to be allowed under any circumstances. The major reason for reverting to the old practice of non-substitution is that the rate of substitution tends to be uneven in a national survey.\(^1\) This lack of uniformity in substitution causes an imbalance in the total sample surveyed across the states. Furthermore, a usual reason for substitution

---

**Table 1:** Total Achieved Sample and Variables in the National Election Studies 1967-2009

<table>
<thead>
<tr>
<th>Year</th>
<th>Achieved Sample Size</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1967</td>
<td>2,287</td>
<td>342</td>
</tr>
<tr>
<td>1971</td>
<td>3,800</td>
<td>255</td>
</tr>
<tr>
<td>1980</td>
<td>3,789</td>
<td>71</td>
</tr>
<tr>
<td>1996</td>
<td>9,614</td>
<td>183</td>
</tr>
<tr>
<td>1998</td>
<td>8,133</td>
<td>44</td>
</tr>
<tr>
<td>1999</td>
<td>9,418</td>
<td>119</td>
</tr>
<tr>
<td>2004</td>
<td>27,189</td>
<td>160</td>
</tr>
<tr>
<td>2009</td>
<td>36,169</td>
<td>280</td>
</tr>
</tbody>
</table>

---

**Table 2:** All-India and State-wise Distribution of Targeted and Achieved Sample, NES 2009

<table>
<thead>
<tr>
<th>States No</th>
<th>AC</th>
<th>PS</th>
<th>Target Sample</th>
<th>Achieved Sample</th>
<th>Achieved Rate (%)</th>
<th>Weighted Share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andhra Pradesh</td>
<td>42</td>
<td>168</td>
<td>3,500</td>
<td>2,508</td>
<td>71.7</td>
<td>6.1</td>
</tr>
<tr>
<td>Arunachal Pradesh</td>
<td>8</td>
<td>32</td>
<td>950</td>
<td>349</td>
<td>36.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Assam</td>
<td>25</td>
<td>100</td>
<td>2,200</td>
<td>1,402</td>
<td>63.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Bihar</td>
<td>40</td>
<td>160</td>
<td>3,200</td>
<td>1,935</td>
<td>60.5</td>
<td>7.7</td>
</tr>
<tr>
<td>Chhattisgarh</td>
<td>15</td>
<td>60</td>
<td>1,200</td>
<td>772</td>
<td>64.3</td>
<td>2.2</td>
</tr>
<tr>
<td>Goa</td>
<td>12</td>
<td>48</td>
<td>1,000</td>
<td>584</td>
<td>58.4</td>
<td>0.1</td>
</tr>
<tr>
<td>Gujarat</td>
<td>26</td>
<td>104</td>
<td>2,200</td>
<td>1,409</td>
<td>64.0</td>
<td>5.1</td>
</tr>
<tr>
<td>Haryana</td>
<td>18</td>
<td>72</td>
<td>1,500</td>
<td>701</td>
<td>46.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Himachal Pradesh</td>
<td>15</td>
<td>60</td>
<td>1,500</td>
<td>837</td>
<td>55.8</td>
<td>0.6</td>
</tr>
<tr>
<td>Jammu and Kashmir</td>
<td>17</td>
<td>68</td>
<td>1,500</td>
<td>919</td>
<td>61.3</td>
<td>0.9</td>
</tr>
<tr>
<td>Jharkhand</td>
<td>19</td>
<td>76</td>
<td>1,600</td>
<td>617</td>
<td>38.6</td>
<td>2.5</td>
</tr>
<tr>
<td>Karnataka</td>
<td>29</td>
<td>116</td>
<td>2,500</td>
<td>2,118</td>
<td>84.7</td>
<td>5.4</td>
</tr>
<tr>
<td>Kerala</td>
<td>20</td>
<td>80</td>
<td>1,600</td>
<td>811</td>
<td>50.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Madhya Pradesh</td>
<td>14</td>
<td>56</td>
<td>1,500</td>
<td>1,616</td>
<td>64.6</td>
<td>5.3</td>
</tr>
<tr>
<td>Maharashtra</td>
<td>48</td>
<td>192</td>
<td>4,000</td>
<td>2,459</td>
<td>61.5</td>
<td>10.2</td>
</tr>
<tr>
<td>Manipur</td>
<td>10</td>
<td>40</td>
<td>950</td>
<td>645</td>
<td>67.9</td>
<td>0.2</td>
</tr>
<tr>
<td>Meghlaya</td>
<td>18</td>
<td>72</td>
<td>1,500</td>
<td>781</td>
<td>52.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Mizoram</td>
<td>18</td>
<td>72</td>
<td>1,500</td>
<td>817</td>
<td>54.5</td>
<td>0.1</td>
</tr>
<tr>
<td>Nagaland</td>
<td>18</td>
<td>72</td>
<td>1,500</td>
<td>787</td>
<td>52.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Orissa</td>
<td>21</td>
<td>84</td>
<td>1,680</td>
<td>1,150</td>
<td>68.5</td>
<td>3.8</td>
</tr>
<tr>
<td>Punjab</td>
<td>20</td>
<td>80</td>
<td>1,600</td>
<td>993</td>
<td>62.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Rajasthan</td>
<td>25</td>
<td>100</td>
<td>2,000</td>
<td>1,259</td>
<td>63.0</td>
<td>5.2</td>
</tr>
<tr>
<td>Sikkim</td>
<td>10</td>
<td>40</td>
<td>950</td>
<td>463</td>
<td>48.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>39</td>
<td>156</td>
<td>3,120</td>
<td>2,629</td>
<td>84.3</td>
<td>5.8</td>
</tr>
<tr>
<td>Tripura</td>
<td>18</td>
<td>72</td>
<td>1,500</td>
<td>998</td>
<td>66.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>80</td>
<td>400</td>
<td>5,200</td>
<td>2,849</td>
<td>54.8</td>
<td>16.3</td>
</tr>
<tr>
<td>Uttarakhand</td>
<td>15</td>
<td>60</td>
<td>1,200</td>
<td>732</td>
<td>61.4</td>
<td>0.8</td>
</tr>
<tr>
<td>West Bengal</td>
<td>42</td>
<td>168</td>
<td>3,500</td>
<td>2,023</td>
<td>57.8</td>
<td>7.4</td>
</tr>
<tr>
<td>Delhi</td>
<td>25</td>
<td>100</td>
<td>2,500</td>
<td>1,001</td>
<td>40.0</td>
<td>1.6</td>
</tr>
<tr>
<td>All India</td>
<td>722</td>
<td>2,808</td>
<td>59,650</td>
<td>36,169</td>
<td>60.6</td>
<td>100.0</td>
</tr>
</tbody>
</table>

PC = parliamentary constituency; AC = assembly constituency; PS = polling station.

Proportion of the total sample achieved refers to the total (unweighted) sample achieved in a state, while weighted share refers to the post-weightage share of the sample from a particular state that duly reflects its share in the national electorate. The information in this table is tentative pending the final cleaning of the data that is currently underway.

(i) Proportion of the total sample achieved refers to the total (unweighted) sample achieved in a state, while weighted share refers to the post-weightage share of the sample from a particular state that duly reflects its share in the national electorate.

(ii) The information on the number of ACs and PSs is subject to the final cleaning of the data that is currently underway.
is to protect against attrition of sample size due to low rate of achievement. But, this can be done by increasing the targeted population without having to resort to substitution. This is what has been done in the 2009 NES. Therefore, even though the allowable rate is lower in the 2009 NES (60.6%) than it was in 2004 (76.9%), where substitution had been allowed, the actual size of achieved sample is much larger in 2009 (36,169, compared to 27,189 in 2004).

Table 2 provides a comprehensive account of the targeted and achieved sample in the NES.

**Interview Schedule**

For the first time in NES history, a “split sample” was introduced. This basically means that rather than one interview schedule being administered to all respondents, five such sets of questionnaires with common questions and background questions, along with certain set specific questions were randomly administered to the respondents. These additional set specific survey modules were on a range of different topics, covering topics such as the economy, security, communalism, democracy and social values. This split sample has provided us with a breadth of questions posed, including details about the contribution to household income by housewives, more nuanced questions on family background and local issues and local politics relevant to each Indian state were also asked. This has provided a rich source of data on opinions regarding the local and national level were introduced. Furthermore, in addition to the five additional themed sets of questions, state-specific questions regarding local issues and local politics relevant to each Indian state were also asked. This has provided a rich source of data on opinions regarding state-level political issues in addition to opinions on national issues.

A pilot study was done in eight states in March 2009 to test the efficacy of the questions and survey instruments to be used in the final survey, but also provide feedback on how the field investigators handle the interviews. The final instruments and the manual for field investigators were revised in the light of this feedback.

It is important to mention that the NES uses a dummy ballot paper to record voting preferences. The respondents were given a (dummy) secret ballot paper designed to look like the list of candidates on the EVM, on which they could mark their choice. These were then placed in a sealed ballot box. This process was followed to ensure that the voters who were interviewed knew their response would be confidential. The introduction of the EVM through the country has meant that voters do not use a ballot paper or a ballot box in the elections any more, yet this procedure of using a “dummy” secret ballot is useful in assuring the respondents of the confidentiality of their response in the interview setting.

**Administration of the NES**

An important component of any data collection exercise is the collective smooth working of a team of investigators, supervisors, coordinators and researchers. Due to the vast coverage network required to conduct a mammoth survey like the NES, a well-trained team was put into place across the country. A team of 39 state coordinators were responsible for coordinating the fieldwork in their respective states. One to three coordinators worked in each state (see Appendix 4, p 202, for a list of coordinators). These coordinators while ensuring regular and accurate data collection in their states also monitored data quality and provided regular transfer of completed questionnaires and entered data to the Lokniti office at Delhi. These coordinators had a team of state supervisors working under them, whose responsibility was to supervise the field investigators (FIs) who conducted the actual interviews. During the NES 2009 fieldwork, a team of two investigators were selected to conduct field investigation for each assembly constituency.

The role of the supervisors was to monitor the work of the FIs, to spot-check some of the selected interviews, and to ensure that the fieldwork conducted met the high standard expected. They also had to ensure that the questionnaires were complete and filled in consistently. These supervisors monitored a total of 1,847 FIs who actually conducted the interviews. The role of the FIs included: locating the

![Table 3](image-url)
voter in the sample, conducting the interview, filling in the schedule for those not interviewed and coding the questionnaire after the completion of the field interviews.

The actual interviews were conducted within a few days of the date of polling in that locality but before the counting day. This meant that the survey too was conducted in five phases to coincide with the five phases of the national elections. The principal reason for conducting the survey before the declaration of results is to ensure that the answers to the survey questions are in no way influenced by the results of the elections. This posed a serious challenge to the survey operations in those constituencies which went to polls in the fifth phase. This included all the constituencies in Tamil Nadu, and some constituencies in West Bengal, Himachal Pradesh, Jammu and Kashmir, Punjab, Uttarakhand and Uttar Pradesh, for it meant that the survey too was conducted in a few days of the date of polling in that constituency but before the counting day. This hands-on experience to potential investigators, to complete the survey left only 72 hours for the survey to be completed.

As FIs form the backbone of any data collection exercise, a rigorous training procedure was put into place for the training and selection of the over 1,800 FIs for the NES. Three Trainers’ Training workshops were held (at Delhi, Guwahati and Hyderabad) for state coordinators and supervisors who would, in turn, train the FIs. Three-day training workshops were held in each state, where the trainers provided hands-on experience to potential FIs. The training workshops were conducted in an interactive format using audio-visuals and with discussions on questionnaires and other survey instruments. Based on mock interviews and a written test, a team of 1,847 FIs were selected for the 2009 NES.

A novel aspect of the 2009 NES was the collection of feedback forms with data on the profiles of the field investigators. This has been an important exercise as this data helps us gauge whether there is any bias in terms of gender, caste, and so on in the team of FIs selected that may have an impact on the completion of interviews or under-representation of certain communities, and which may be corrected in future surveys. Of the feedback forms received, the selection of the FIs has been very well spread across the country, and rather than any one state providing a bulk of the interviewers we have had a good representation of our network across the country. Also, 31% of the FIs have worked previously with the cso on a survey, and 15% have worked on other surveys, which taken together suggest that a large proportion of the FIs had prior survey and fieldwork experience. There is, however, seen to be a big over-representation of male FIs, while only 24% of the FIs were women. This is an aspect that needs to be borne in mind for future surveys, as increasing the proportion of female FIs may help increase the rate of completion of our surveys by women (which is currently slightly lower than that for men). In terms of caste and community, the FIs are fairly representative of the Indian population.

**Table 4: Religion and Community Profile by State – NES 2009 vs Census of India 2001 (in %)**

out the possibility of a selection bias in the sampling.

This seems to be the case in terms of community representativeness as well, as on average the all India figures very closely match that of the last census. Arunachal Pradesh stands out as the only state with a substantial (15 percentage point) under-representation of Hindus when compared to the census figures, matched partly by the over-representation (10 percentage points) of the Christians in the state sample. Further, we see a slight over-representation of dalits across the states (with a couple of minor exceptions, such as Himachal Pradesh and Manipur). With regard to the adivasi community, Sikkim and Arunachal Pradesh are two states with a substantial over-representation of this community (adivasis account for more than 21 percentage points over the census figures for these states) in the survey data.

An innovative feature of NES 2009 was that information was recorded on the age, sex, religion, caste of the person not interviewed and of course reasons for non-interview. A study of this data helps ensure that those who did not take part in the survey were not substantially different from those who actually did. This data was collated and analysed to study any obvious patterns in the non-response. Among the many reasons for non-interview (Table 5), the most common was being “temporarily away” (42%). More importantly, about 32% of those not interviewed were people with incomplete, missing or outdated information in the electoral rolls (including 5% who were dead, 12% with unknown addresses and 15% who are permanently away). Only about 6% of the sample population actually refused to answer the survey questions.

In terms of their demographic profile, the gender profile of those not interviewed is nearly identical to that of the final NES (53.4% men and 46.6% women). The age profile of the respondents too is broadly similar among the non-interviewed as it is for the interviewed. Further, in terms of caste as well, the non-interviewed have a very similar profile to the interviewed. The one exception is with regard to the upper castes, where in the non-interviewed sample just over 2.5% more respondents were from the upper caste compared to the respondents who were actually interviewed. The rest of the caste profile is broadly similar to that of the interviewed respondents. Finally, in terms of their religious profile we see a slightly greater difference between the interviewed and non-interviewed sample primarily with regard to Hindus and respondents from other religions. The non-interviewed had a slight under-sample of Hindus with a subsequent over-sample of respondents from other religions. But, as the final NES 2009 sample is quite similar to the 2001 Census with respect to religion, this slight discrepancy in the non-interviewed sample is not of great concern. Given the broad similarity in the demographic profile of both the interviewed and the non-interviewed sample, we can be fairly confident in not having any specific selection bias due to non-response in the final NES dataset.

In party-political terms, the NES captures the picture fairly well at the national level. Table 6 presents data on the state-wise comparison of actual vote share in the Lok Sabha elections reported by the Election Commission and the reported vote share in the 2009 NES. States such as Himachal Pradesh, Karnataka and Mizoram display a very accurate assessment of the vote share for the major parties. Whereas, we see that in Assam, Haryana, Nagaland, Sikkim and Jharkhand there are greater discrepancies from the official figures.

### Data Analysis

The final step of the data collection process includes the entering, rigorous checking and cleaning of the data. This was done by a team of experienced technicians at the CSDS data unit in Delhi (see Appendix 5, p 202). The cleaning and cross-checking of data is still under process and the Data Unit has taken special efforts to make the data available for the analyses in the papers of this special issue.

The inaccuracies and inconsistencies of the demographic profile and the actual vote share are corrected for by using the procedure of weightage. Following generally accepted statistical norms, data analysis based on the NES 2009 often employs two types of weights depending on the statistical analysis involved: the first is the voter turnout weight and the second is the vote share weight. Furthermore, the analysis at the national level is done by weighing the sample by the states’ share in national electorate.

---

**Table 6: Actual Vote Share and Reported Vote Share in NES 2009 by State (in %)**

<table>
<thead>
<tr>
<th>States</th>
<th>UPA Actual Survey</th>
<th>UPA Actual Survey</th>
<th>NDA Actual Survey</th>
<th>NDA Actual Survey</th>
<th>Other Major Parties/Alliances (More than 5% Votes)</th>
<th>Other Major Parties/Alliances (More than 5% Votes)</th>
<th>Others Actual Survey</th>
<th>Others Actual Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andhra Pradesh</td>
<td>39.0</td>
<td>44.5</td>
<td>3.7</td>
<td>1.6</td>
<td>TDP+ PRP</td>
<td>33.6</td>
<td>38.5</td>
<td>15.7</td>
</tr>
<tr>
<td>Arunachal Pradesh</td>
<td>51.1</td>
<td>42.3</td>
<td>37.1</td>
<td>49.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Assam</td>
<td>34.9</td>
<td>46.6</td>
<td>30.8</td>
<td>33.1</td>
<td>AUDF</td>
<td>16.1</td>
<td>12.5</td>
<td>12.8</td>
</tr>
<tr>
<td>BPF</td>
<td>5.4</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bihar</td>
<td>10.3</td>
<td>13.2</td>
<td>38.0</td>
<td>43.6</td>
<td>RJ+</td>
<td>25.8</td>
<td>31.7</td>
<td>25.9</td>
</tr>
<tr>
<td>Chhattisgarh</td>
<td>37.3</td>
<td>45.3</td>
<td>45.3</td>
<td>43.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Delhi</td>
<td>57.1</td>
<td>58.4</td>
<td>35.5</td>
<td>31.8</td>
<td>BSP</td>
<td>5.3</td>
<td>7.6</td>
<td>2.1</td>
</tr>
<tr>
<td>Goa</td>
<td>45.9</td>
<td>41.2</td>
<td>44.8</td>
<td>55.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gujarat</td>
<td>43.5</td>
<td>47.7</td>
<td>46.6</td>
<td>49.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Haryana</td>
<td>41.8</td>
<td>53.4</td>
<td>27.9</td>
<td>28.2</td>
<td>BSP</td>
<td>15.8</td>
<td>10.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Himachal Pradesh</td>
<td>45.6</td>
<td>47.9</td>
<td>49.5</td>
<td>50.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Jammu and Kashmir</td>
<td>43.8</td>
<td>50.9</td>
<td>16.6</td>
<td>20.4</td>
<td>PDP</td>
<td>20.1</td>
<td>22.8</td>
<td>17.5</td>
</tr>
<tr>
<td>Jharkhand</td>
<td>26.6</td>
<td>31.7</td>
<td>28.7</td>
<td>39.6</td>
<td>JVM</td>
<td>10.5</td>
<td>9.1</td>
<td>28.9</td>
</tr>
<tr>
<td>Karnataka</td>
<td>37.7</td>
<td>37.7</td>
<td>41.6</td>
<td>42.4</td>
<td>JD (S)</td>
<td>13.6</td>
<td>15.7</td>
<td>7.1</td>
</tr>
<tr>
<td>Kerala</td>
<td>47.7</td>
<td>54.5</td>
<td>41.6</td>
<td>42.4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Madhya Pradesh</td>
<td>40.1</td>
<td>38.4</td>
<td>43.5</td>
<td>49.8</td>
<td>BSP</td>
<td>5.9</td>
<td>8.2</td>
<td>10.5</td>
</tr>
<tr>
<td>Maharashtra</td>
<td>40.2</td>
<td>52.9</td>
<td>35.2</td>
<td>30.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Manipur</td>
<td>43.0</td>
<td>44.2</td>
<td>9.5</td>
<td>8.3</td>
<td>CPI</td>
<td>14.9</td>
<td>14.4</td>
<td>19.0</td>
</tr>
<tr>
<td>Meghalaya</td>
<td>44.8</td>
<td>54.9</td>
<td>-</td>
<td>-</td>
<td>NCP</td>
<td>15.1</td>
<td>15.9</td>
<td>22.0</td>
</tr>
<tr>
<td>Mizoram</td>
<td>65.6</td>
<td>67.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Nagaland</td>
<td>29.4</td>
<td>40.7</td>
<td>-</td>
<td>-</td>
<td>NPP</td>
<td>70.0</td>
<td>54.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Orissa</td>
<td>32.5</td>
<td>39.4</td>
<td>16.9</td>
<td>23.1</td>
<td>BJD+</td>
<td>41.7</td>
<td>36.9</td>
<td>8.9</td>
</tr>
<tr>
<td>Punjab</td>
<td>45.2</td>
<td>49.8</td>
<td>43.9</td>
<td>45.1</td>
<td>BSP</td>
<td>5.7</td>
<td>3.3</td>
<td>5.2</td>
</tr>
<tr>
<td>Rajasthan</td>
<td>47.2</td>
<td>53.9</td>
<td>36.6</td>
<td>40.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sikkim</td>
<td>29.6</td>
<td>16.6</td>
<td>1.8</td>
<td>4.4</td>
<td>SDF</td>
<td>63.3</td>
<td>53.7</td>
<td>5.3</td>
</tr>
<tr>
<td>Tamil Nadu</td>
<td>42.5</td>
<td>48.5</td>
<td>2.3</td>
<td>1.5</td>
<td>ADMK+ DMK</td>
<td>37.4</td>
<td>32.8</td>
<td>6.7</td>
</tr>
<tr>
<td>Tripura</td>
<td>30.7</td>
<td>28.0</td>
<td>3.4</td>
<td>1.8</td>
<td>CP</td>
<td>61.7</td>
<td>67.4</td>
<td>4.2</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>18.2</td>
<td>19.8</td>
<td>20.8</td>
<td>23.0</td>
<td>SP</td>
<td>23.3</td>
<td>23.0</td>
<td>10.3</td>
</tr>
<tr>
<td>Uttar Pradesh</td>
<td>43.3</td>
<td>42.2</td>
<td>34.0</td>
<td>37.7</td>
<td>BSP</td>
<td>15.3</td>
<td>14.5</td>
<td>7.4</td>
</tr>
<tr>
<td>West Bengal</td>
<td>44.7</td>
<td>39.4</td>
<td>6.1</td>
<td>5.2</td>
<td>LF</td>
<td>43.4</td>
<td>52.6</td>
<td>5.8</td>
</tr>
<tr>
<td>All India</td>
<td>36.4</td>
<td>38.8</td>
<td>34.1</td>
<td>26.6</td>
<td>LF</td>
<td>7.6</td>
<td>6.7</td>
<td>10.6</td>
</tr>
</tbody>
</table>

---

**EPW Weekly September 26, 2009 VOL XLIV NO 39 ECONOMICS & POLITICAL WEEKLY**
**Technical Note**

*N* refers to the sample size of the data. The *N* for the entire NES 2009 dataset is 36,169.1 However, due to two main reasons the *N* reported in the papers in this issue may vary from this total *N*. First, the *N* reported varies according to the weights that have been employed (discussed below) for that particular analysis. For example data weighted by vote share or by turnout will have slightly different *Ns*. Second, the difference in *Ns* may be because of the removal of missing data. As any data that is missing with regard to the variables under study, has been removed from the analysis, the reported sample size for that particular analysis will differ from the full sample *N*. For example, if we present the analysis for party-vote, we would remove from our analysis anyone who did not vote in the elections thus reducing the *N*.

**Weights:** The NES 2009 data is collected through a probability sample at the state level which is added up to give the national sample. However, in order to make the data representative at the national level it is weighted to ensure that the proportion of each state reflects its share in the national electorate and so that we do not over-represent the smaller states. This weight is arrived at by taking the total official electorate proportion of the state and dividing it by the state’s proportion in the NES 2009. The figures for the weighted share of each state are presented in the last column of Table 2. In addition to the weight for state proportion two other weights are applied depending on the analysis involved: the first is the weight for voter turnout and the second is the weight for vote share.

_turnout weight:_ The reported turnout in the NES 2009 is 83.6%, while the actual turnout according to the ECI was 58.4%. As the over-reporting of turnout is a major issue not only in Indian election surveys, but in national elections surveys around the world, we apply weights for actual turnout (figures from the ECI) to the NES 2009 dataset. This is done by using multi-weights as seen in the following formula:

\[
\text{State electorate proportion weight} \times \text{(actual turnout figures reported by the ECI)}
\]

**Recoded Variables:** Two primary recoded variables used are caste/community and class. We now describe these variables along with their coding and syntax:

**Caste/Community:** The community variable used in the analyses (scaste), unless otherwise indicated, combines information from two questions in the NES: on religion (z8 in the NES 2009 questionnaire) and on the respondent’s caste/jat/biradri/tribe name (z7 in the NES 2009 questionnaire). The final categories under this variable are:

I) Hindu upper castes
II) Hindu peasant proprietors
III) Hindu upper OBC
IV) Hindu lower OBC
V) Hindu dalit
VI) Hindu adivasi
VII) Muslims
VII) Others

In this classification, caste categories have been treated as essentially Hindu to contrast them to the Muslims and other religious communities (see syntax below). This implies that in the community coded variable, dalits and adivasis are restricted to those that reported their religion as Hindu and are in effect Hindu dalits. Hindu adivasis 2 For the caste categories themselves, for people with regard to dividing the OBCs into upper and lower, a subjective classification was followed. This classification was based on an analysis of the socio-politico-economic status that each caste enjoys in their state. The caste coding is provided in the syntax below. For analysis in states in which there were small sample sizes, the eightfold caste/community categorisation described above was reduced by collapsing certain categories appropriately, for example by merging the upper and lower OBCs.

**Syntax for community (scaste):**

```
Compute Scaste=0.
If (Z8=1 and (Z7>=1 and Z7<=99)) Scaste=1.
If (Scaste=0 and Z8=1 and (Z7>=100 and Z7=199)) Scaste=2.
If (Scaste=0 and Z8=1 and (Z7>=200 and Z7=299)) Scaste=3.
If (Scaste=0 and Z8=1 and (Z7>=300 and Z7=499)) Scaste=4.
If (Scaste=0 and Z8=1 and (Z7>=500 and Z7=699)) Scaste=5.
If (Scaste=0 and Z8=1 and (Z7>=700 and Z7=699)) Scaste=6.
If (Scaste=0 and Z8=2) Scaste=7.
```

**Value label Scaste 1 '1: Upper caste' 2: Peasant Proprietors' 3 '3: Upper OBC' 4 'Lower OBC' 5 'Dalit 6 '6: ST' 7 'Muslims' 8 '8: Others'.**

**Variable label Scaste 'Caste community'.**

**Class:** The class variable used, unless otherwise stated, is a combination of information on both monthly household income (z18 in the NES 2009 questionnaire) and on household assets (z16 in the NES 2009 questionnaire). A five class schema ranging from Rich to Very Poor was created based on different income levels and asset ownership. For example, a respondent with household income of less than Rs 1,000 a month and who owns none of the assets listed in the dataset was assigned to the Very Poor class category. The household assets used to create this class schema included the following (with the particular questionnaire code in brackets): bicycle (z16a), LPG (z16b), Number of Telephones (z16c), Electric fan/cooler (z16d), Black and White television (z16e), Colour Television (z16f), Cable connection (z16g), Number of Scooters/motorcycles/mopeds (z16h), Number of Cars/Jeeps/Vans (z16i), Tractor (z16j) and Fridge (z16k). The categorisation of which particular asset was included in each class is shown in the syntax below.

A correction was made to the schema regarding the ownership of LPG gas in the household. As the ownership of LPG can be considered to be fairly widespread in urban areas but not in rural areas, the class schema assigned respondents in rural areas who owned LPG gas to the Rich class category. The monthly household income categories used were: Income 1: less than 1000; Income 2: 1000 to 2000; Income 3: 2001 to 3000; Income 4: 3001 to 4000; Income 5: 4001 to 5000; Income 6: 5001 to 10000; Income 7: 10001 to 20000; Income 8: 20001 to highest income.

**Syntact for class:**

```
Recode z18 (0 thru 1000=1) (1001 thru 2000=2) (2001 thru 3000=3) (3001 thru 4000=4) (4001 thru 5000=5) (5001 thru 10000=6) (10001 thru 20000=7) (20001 thru Highest=8) (ELSE=SYSMIS) into income.
```

**Variable labels income 'income'.**

**Execute .**

**Compute Lpg=0.**

```
If (Loca=1 and Z16b=1) Lpg=1.
If (Lpg=0) Lpg=0.
```

**Format Lpg (F1.0).**

**Compute Class=0.**

```
If (Z16c or Z16d or Z16e or Z16f or Z16g or Z16h) Class=1.
If (Class=0 and (Z16f=1 and Z16h=1 and Z16i=1 and Z16k=1 and Lpg=1)) Class=1.
```

**Syntax for class:**

```
If (Class=0 and Income=8) Class=1.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1)) Class=2.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16j=1)) Class=2.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16j=1 and Z16k=1)) Class=2.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16k=1 and Z16j=1)) Class=2.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16k=1 and Z16j=1 and Z16i=1)) Class=3.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16i=1 and Z16k=1 and Z16j=1)) Class=3.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16i=1 and Z16j=1 and Z16k=1 and Z16l=1)) Class=3.
If (Class=0 and (Z16c=1 and Z16f=1 and Z16h=1 and Z16i=1 and Z16j=1 and Z16k=1 and Z16l=1 and Z16m=1)) Class=3.
```

**Execute .**

**Value label Class 1 '1: Rich' 2 'Middle' 3 'Lower' 4 'Poor' 5 'Very poor'.**

**Variable label Class 'Economic class by assets and income' execute.**

---

1. This total is tentative pending the final cleaning of the data that is currently underway.

2. Any dalit or adivasi from a different religion, for example 54th dalits or Christian adivasis, are placed in the other category according to this coding.

3. For papers on these communities themselves, e.g. the papers on dalit and adivasi electoral participation, this caste community categorisation was not used, but instead the question on caste group (27a in the NES questionnaire) was used to allow for other religions among the dalits and adivasis themselves and not restricting them to Hindus only.
Funding
The NES 2009 has the unique distinction of being a collaborative research project shared by over 35 scholars from across the country. This election study was made possible by the financial support extended by two premier organisations of higher education in India. The University Grants Commission (UGC) granted Major Research Projects to 22 scholars of the Lokniti network from different universities and colleges in the country. This financial support received from UGC made it possible for the scholars concerned to be part of the collective programme of the NES and conduct the election study in their respective states. Financial support to Lokniti for NES 2009 was also provided by the Indian Council of Social Science Research (ICSSR) for designing the research instruments, workshop for trainers from the Lokniti network members and for workshops for FIs all over the country. One thousand eight hundred and forty seven field investigators were trained in 39 workshops held across the country for this purpose.2

Apart from this support, the survey was also financially supported by The Hindu newspaper and the CNN-IBN television news channel. The preliminary results of our survey were distributed via these media organisations both during and after the elections.

Conclusions
Lokniti aims at developing an understanding of India’s electoral politics in particular and democratic political process in general. The goal is to ground political analyses firmly in political science in India by ensuring that data thus collected is used for writings on Indian politics and society. After the Lokniti team has conducted the interviews in each PS to ensure a faster and smoother collection of data.

The NES data is available to students of Indian politics and society. After the Lokniti team has used the data, it will be made available to the academic community at a cost rate. There is a special scheme to encourage students based in Indian universities to download this data. For further details visit our website at www.lokniti.org.

REFERENCES


Appendix 1: Lokniti Research Team
Sanjay Kumar of ccsrs, Delhi coordinated this All India Survey. The research team at Lokniti that coordinated and analysed the survey comprised Sanjeet Alam, Banasmita Bora, Vanita Lehav Falcao, Vikas Gautam, Dhananjay Joshi, Navpriet Kaur, Praveen Rai, Ram Narayan Ram, Dhananjay Kumar Singh, Divya Vaid and Rahul Verma.

Appendix 2: Interview Schedule Designing Committee
Rajeshwari Deshpande, Dhananjai Joshi, K K Kailash, Sanjay Kumar, Sanjay Lodha, Suhas Palshikar, Praveen Rai, Sandeep Shastri, D L Sheth, V B Singh, K C Suri, and Yogendra Yadav

Appendix 3: List of Languages into Which the Interview Schedules for NES 2009 Were Translated

Appendix 4: List of State Coordinators

State Name of the Coordinator
Andhara Pradesh P Narasimha Rao
Arunachal Pradesh Nani Bath
Assam Sandhya Goswami
Bihar Rakesh Ranjan
Chhattisgarh Anupama Saxena
Delhi Biswajit Mohanty
Goa Maria do Ceu Rodrigues
Gujarat Mahashweta Jini and Bhanu Parmar
Himachal Pradesh Ramesh Chauhan
Haryana Harish Kumar and Kushal Pal
Jammu and Kashmir Ellora Puri and Gul Mohammad Wani
Jharkhand Harishwar Dayal and B K Sinha
Karnataka Veena Devi and Padmavathi B S
Kerala Sajaj Ibrahaim
Madhya Pradesh Ram Shankar Dubey and yatindra Singh sisodia
Maharashtra Nitin Birmal
Manipur Senjam Mangi Singh
Meghalaya R K Satapathy
Mizoram Lallian Chhunga
Nagaland Amongla Jamir
Orissa Surya Narayan Misra
Punjab Ashutosh Kumar and Jagroop Singh Sekhon
Rajasthan Sanjay Lodha
Sikkim Mukund Gin
Tamil Nadu Gundapuneni Koteswara Prasad
Tripura Sukhendu Debbarma
Uttar Pradesh Mirza Asmer Beg, Sudhir Kumar and A K Verma
Uttarakhand Annpurna Nautiyal
West Bengal Supono Basu

Appendix 5: Data Unit Team
Himanshu Bhattacharya, K A Q A Hilal and Kanchan Mallhotra of the ccsrs Data Unit.