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-Palshikar 2003 and 2009; Special Issues of the EPW, 1999 and 2004; *JISPE* 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.[1] 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

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

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

of questions on various socio-politico-economic themes were asked, the collection of data on the non-interviewed sample to monitor the representativeness of the NES, the introduction of a rigorous back-checking procedure to establish the robustness of the data-collection exercise, the rigorous training of field investigators as well as collection of information on these FIs to ensure better representativeness in the future, are some of the recent methodological innovations. A team of researchers coordinated this entire effort from December 2008 onwards (see Appendix 1, p 202).

Sampling

The sampling scheme adopted in NES 2009 was a repetition of the practice followed in NES 2004 with a few modifications. As in 2004, the sample was drawn by using four-stage stratified random sampling. This procedure ensures that the selected sample is fully representative of the cross-section of voters in the country. A major methodological departure in NES 2004, when compared to the previous NES surveys, had been to abandon the earlier practice of national self-weighted probability sample in favour of a probability sample at the state level which is added up and weighted to arrive at a national representative sample. The same practice was followed this time, except that the target sample was enhanced for some states, especially the smaller ones. Thus the NES has been designed to be representative not just at the national level, but also at the state level. Table 2 provides data on sampling at each of these stages.

The first stage of the stratified random sample has become, for all practical purposes, a census rather than a sample. In all, 536 PCs in the 28 states of the country and union territory of Delhi were selected for coverage, out of the total 543 PCs in India. Ladakh in Jammu and Kashmir and the PCs falling in the union territories of Pondicherry, Chandigarh,[2] Lakshadweep, Andaman and Nicobar Islands, Dadra and Nagar Haveli, and Daman and Diu

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 in the big states 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

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

Sl No	States	AC	PS	Target Sample	Achieved Sample	Achieval Rate (%)	Weighed Share (%)
1	Andhra Pradesh	42	168	3,500	2,508	71.7	8.1
2	Arunachal Pradesh	8	32	950	349	36.7	0.1
3	Assam	25	100	2,200	1,402	63.7	2.5
4	Bihar	40	160	3,200	1,935	60.5	7.7
5	Chhattisgarh	15	60	1,200	772	64.3	2.2
6	Goa	12	48	1,000	584	58.4	0.1
7	Gujarat	26	104	2,200	1,409	64.0	5.1
8	Haryana	18	72	1,500	701	46.7	1.7
9	Himachal Pradesh	15	60	1,500	837	55.8	0.6
10	Jammu and Kashmir	17	68	1,500	919	61.3	0.9
11	Jharkhand	19	76	1,600	617	38.6	2.5
12	Karnataka	29	116	2,500	2,118	84.7	5.4
13	Kerala	20	80	1,600	811	50.7	3.1
14	Madhya Pradesh	29	116	2,500	1,616	64.6	5.3
15	Maharashtra	48	192	4,000	2,459	61.5	10.2
16	Manipur	10	40	950	645	67.9	0.2
17	Meghalaya	18	72	1,500	781	52.1	0.2
18	Mizoram	18	72	1,500	817	54.5	0.1
19	Nagaland	18	72	1,500	787	52.5	0.2
20	Orissa	21	84	1,680	1,150	68.5	3.8
21	Punjab	20	80	1,600	993	62.1	2.4
22	Rajasthan	25	100	2,000	1,259	63.0	5.2
23	Sikkim	10	40	950	463	48.7	0.1
24	Tamil Nadu	39	156	3,120	2,629	84.3	5.8
25	Tripura	18	72	1,500	998	66.5	0.3
26	Uttar Pradesh	80	240	5,200	2,849	54.8	16.3
27	Uttarakhand	15	60	1,200	737	61.4	0.8
28	West Bengal	42	168	3,500	2,023	57.8	7.4
29	Delhi	25	100	2,500	1,001	40.0	1.6
	All India	722	2,808	59,650	36,169	60.6	100.0

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.

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 PSs 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 to 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 of 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 PS 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 drawn 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.[3] This lack of uniformity in substitution causes an imbalance in the total sample surveyed across the states. Furthermore, a usual reason for substitution

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 achieval 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).[4] 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 themes, covering topics such as the economy, security, communalism, democracy and social values. This split sample has provided us with a veritable treasure trove of data on opinions on issues beyond those dealing with political and electoral processes. The five sets of questionnaires randomly administered to the sampled voters were carefully designed by a committee of social scientists (see Appendix 2, p 202) and tested during the March Pilot Survey.

All questions were carefully phrased to avoid the interviewee being led to give any particular answer. The questionnaire was translated into 20 languages (see Appendix 3, p 202). The translation process was carefully monitored by a group of experts, so that a question in one state did not have a different meaning in another.[5]

As in previous NES surveys, in the current one as well, detailed information about the background socio-demographics and the household was collected. However, the 2009 NES stands out as having an even greater breadth of questions posed, including details on the contribution to household income by housewives, more nuanced questions on family income and more specific categories of jati/caste sub-group at the state rather than at the 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 to be used for the main NES. A total of 466 voters randomly sampled from the voter list were interviewed in these eight states. The list of survey instruments tested in these pilot surveys included the sampling frame, respondent list, ballot box, dummy Electronic Voting Machine (EVM) slips and questionnaires. These pilot surveys provide a dual benefit of not only testing 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

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

Table 3: Gender and Locality Profile by State – NES 2009 vs Census of India 2001 (in %)

Sl No	States	Men		Women		Rural		Urban	
		Census	Survey	Census	Survey	Census	Survey	Census	Survey
1	Andhra Pradesh	50.6	47.5	49.5	52.5	72.7	66.6	27.3	33.4
2	Arunachal Pradesh	52.8	47.9	47.2	52.1	79.2	83.7	20.8	16.3
3	Assam	51.7	52.3	48.3	47.7	87.1	80.7	12.9	19.3
4	Bihar	52.1	51.3	47.9	48.7	89.5	86.2	10.5	13.8
5	Chhattisgarh	50.3	52.2	49.7	47.8	79.9	86.4	20.1	13.6
6	Goa	51.0	50.2	49.0	49.8	50.2	92.1	49.8	7.9
7	Gujarat	52.1	54.3	47.9	45.7	62.6	61.2	37.4	38.8
8	Haryana	53.7	53.9	46.3	46.1	71.1	82.5	28.9	17.5
9	Himachal Pradesh	50.8	48.9	49.2	51.1	90.2	97.7	9.8	2.3
10	Jammu and Kashmir	52.9	55.2	47.2	44.6	75.2	76.7	24.8	23.3
11	Jharkhand	51.5	60.4	48.5	39.6	77.8	78.8	22.2	21.2
12	Karnataka	50.9	51.5	49.1	48.5	66.0	68.4	34.0	31.6
13	Kerala	48.6	43.6	51.4	56.4	74.0	81.3	26.0	18.7
14	Madhya Pradesh	52.1	55.3	47.9	44.7	73.5	77.5	26.5	22.5
15	Maharashtra	52.0	55.7	48.0	44.3	57.6	47.5	42.4	52.5
16	Manipur	50.6	51.9	49.4	48.1	73.4	74.3	26.6	25.7
17	Meghalaya	50.7	57.2	49.3	42.8	80.4	77.5	19.6	22.5
18	Mizoram	51.7	49.4	48.3	50.6	50.4	58.3	49.6	41.7
19	Nagaland	52.6	50.8	47.4	49.2	82.8	71.7	17.2	28.3
20	Orissa	50.7	55.6	49.3	44.4	85.0	74.8	15.0	25.2
21	Punjab	53.3	52.3	46.7	47.7	66.1	69.1	33.9	30.9
22	Rajasthan	52.1	55.4	47.9	44.6	76.6	77.4	23.4	22.6
23	Sikkim	53.3	56.2	46.7	43.8	88.9	83.6	11.1	16.4
24	Tamil Nadu	50.3	50.8	49.7	49.2	56.0	54.3	44.0	45.7
25	Tripura	51.3	58.1	48.7	41.9	82.9	70.2	17.1	29.8
26	Uttar Pradesh	52.7	56.8	47.3	43.2	79.2	81.3	20.8	18.7
27	Uttarakhand	51.0	52.5	49.0	47.5	74.3	76.3	25.7	23.7
28	West Bengal	51.7	52.6	48.3	47.4	72.0	76.4	28.0	23.6
29	Delhi	54.9	55.7	45.1	44.3	6.8	17.3	93.2	82.7
	All India	51.7	53.5	48.3	46.5	72.2	71.8	27.8	28.2

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

were selected to conduct field investigation for each assembly constituency.[6]

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

S. No	STATES	Hindus		Other Major Religions			Dalit		Adivasi	
		Census	Survey	Religion	Census	Survey	Census	Survey	Census	Survey
1	Andhra Pradesh	89.0	81.5	Muslims	9.2	10.6	16.2	18.7	6.6	6.3
2	Arunachal Pradesh	34.6	19.8	Christians	18.7	29.5	0.6	3.2	64.2	84.8
3	Assam	64.9	71.5	Muslims	30.9	23.2	6.9	12.1	12.4	15.3
4	Bihar	83.2	83.2	Muslims	16.5	15.3	15.7	19.4	0.9	1.2
5	Chhattisgarh	94.7	93.7	Muslims	2.0	1.8	11.6	13.8	31.8	40.8
6	Goa	65.8	71.9	Christians	26.7	23.3	1.8	6.4	0.0	8.2
7	Gujarat	89.1	86.2	Muslims	9.1	10.1	7.1	8.0	14.8	16.2
8	Haryana	88.2	94.4	Muslims	5.8	0.9	19.3	21.7	0.0	0.4
9	Himachal Pradesh	95.4	92.7	Muslims	2.0	5.0	24.7	20.8	4.0	5.3
10	Jammu and Kashmir	29.6	29.4	Muslims	67.0	64.2	7.6	8.0	10.9	4.0
11	Jharkhand	68.6	64.3	Muslims	13.9	6.9	11.8	11.9	26.3	24.2
12	Karnataka	83.9	84.6	Muslims	12.2	11.5	16.2	20.1	6.6	10.0
13	Kerala	56.2	56.7	Muslims	24.7	22.3	9.8	11.1	1.1	3.9
14	Madhya Pradesh	91.2	89.9	Muslims	6.4	3.8	15.2	20.3	20.3	19.2
15	Maharashtra	80.4	79.2	Muslims	10.6	7.6	10.2	16.2	8.9	11.8
16	Manipur	46.0	56.7	Christians	34.0	30.7	2.8	0.6	34.2	32.3
17	Meghalaya	13.3	9.7	Christians	70.3	66.2	0.5	3.6	85.9	83.5
18	Mizoram	3.6	3.8	Christians	87.0	88.6	0.0	1.7	94.5	97.0
19	Nagaland	7.7	1.8	Christians	90.0	96.8	0.0	1.0	89.1	92.0
20	Orissa	94.4	97.8	Muslims	2.1	1.5	16.5	15.9	22.1	18.0
21	Punjab	36.9	30.9	Sikh	59.9	67.3	28.9	29.7	0.0	1.1
22	Rajasthan	88.8	89.6	Muslims	8.5	8.2	17.2	22.4	12.6	11.4
23	Sikkim	60.9	60.0	Buddhist	28.1	26.8	5.0	7.2	20.6	41.0
24	Tamil Nadu	88.1	85.2	Muslims	5.6	7.8	19.0	19.7	1.0	1.6
25	Tripura	85.6	78.0	Muslims	8.0	6.2	17.4	27.6	31.1	33.5
26	Uttar Pradesh	80.6	78.0	Muslims	18.5	18.7	21.1	22.7	0.1	1.9
27	Uttarakhand	85.0	86.1	Muslims	11.9	9.2	17.9	21.1	3.0	3.7
28	West Bengal	72.5	69.8	Muslims	25.3	28.1	23.0	24.2	5.5	6.4
29	Delhi	82.0	80.9	Muslims	11.7	10.6	16.9	16.4	0.0	1.1
	All India	80.5	79.1	Muslims	13.4	12.8	16.2	18.8	8.2	9.2

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 left only 72 hours for the survey to be completed. Special arrangements were made, including additional investigators, to complete the survey before the deadline.

A new feature of NES 2009 has been the introduction of a rigorous back-checking procedure after the completion of the whole survey. Under this, a selection of 5% of the respondents who completed the original questionnaire were approached in order to check that their responses to the original questionnaire were correctly coded in our database. This procedure further validates the robustness of our data collection exercise and ensures that our data meet the highest standards of international survey research.

Training Workshops

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 CSDS 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 overrepresentation 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.

Sample Profile

Of the 59,650 persons approached for interviews, 36,169 interviews could be completed. The sample profile shows that the persons interviewed were broadly reflective of the Indian population, in terms of the country's general demographic profile. A study of the sample profile (Table 3, p 198 and Table 4) highlights the representativeness of the NES when compared to the last Census of 2001. On average, in Table 3, there is a small (less than 2 percentage point) over-representation of men compared to the census, with a subsequent under-representation of women in the survey. In terms of the rural-urban split of the sample, the NES figures are nearly identical to the census figures at the All India level. However, as with any average figure, this hides some discrepancies at the state level. Some states such as Nagaland, Tripura, Orissa and Maharashtra have a lower sample from rural areas as compared to the census figures for these states, but this is balanced by states such as Goa, Haryana and Himachal Pradesh, to name a few, where we have an over-sampling of the rural population. Despite this slight mismatch between the census and survey figures, no particular pattern in the over or under-reporting of the rural/urban areas emerges, thus ruling

Table 5: Reason for Non-Interview

Reasons	%
Temporarily away	42
Permanently away	15
Address not known	12
Failed to contact	9
Dead	5
Ill	4
Refused to answer	6
Others reasons	7

Sample size is 17,528 non-interviewed respondents.

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 collected with the use of the "dummy" ballot box. The UPA figures in particular are quite close to the actual official figures. 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 %)

Sl No	States	UPA Actual	UPA Survey	NDA Actual	NDA Survey	Other Party	Other Actual	Other Survey	Others Actual	Others Survey
1	Andhra Pradesh	39.0	44.5	3.7	1.6	TDP+ PRP	33.6 15.7	38.5 14.4	8.0	1.0
2	Arunachal Pradesh	51.1	42.3	37.1	35.2	-	-	-	11.8	22.5
3	Assam	34.9	46.6	30.8	33.1	AUDF BPF	16.1 5.4	12.5 0.0	12.8	7.8
4	Bihar	10.3	13.2	38.0	43.6	RJD+	25.8	31.7	25.9	11.5
5	Chhattisgarh	37.3	45.3	45.0	45.3	-	-	-	17.7	9.4
6	Delhi	57.1	58.4	35.5	31.8	BSP	5.3	7.6	2.1	2.2
7	Goa	45.9	41.2	44.8	55.6	-	-	-	9.3	3.2
8	Gujarat	43.5	47.7	46.6	49.1	-	-	-	9.9	3.2
9	Haryana	41.8	53.4	27.9	28.2	BSP HJC	15.8 10.1	10.4 6.4	4.4	1.6
10	Himachal Pradesh	45.6	47.9	49.5	50.7	-	-	-	4.9	1.4
11	Jammu and Kashmir	43.8	50.9	18.6	20.4	PDP	20.1	22.8	17.5	5.9
12	Jharkhand	26.6	31.7	28.7	39.6	JVM RJD	10.5 5.3	9.1 4.8	28.9	14.8
13	Karnataka	37.7	37.7	41.6	42.4	JD (S)	13.6	15.7	7.1	4.2
14	Kerala	47.7	47.2	6.3	3.5	LF	41.9	47.5	4.1	1.8
15	Madhya Pradesh	40.1	38.4	43.5	49.8	BSP	5.9	8.2	10.5	3.6
16	Maharashtra	40.2	52.9	35.2	30.0	NCP	6.0	3.0	24.6	17.1
17	Manipur	43.0	44.2	9.5	8.3	CPI MPP	14.9 7.6	14.4 11.5	19.0	18.6
18	Meghalaya	44.8	54.9	-	-	NCP UDP	18.1 15.1	15.9 9.8	22.0	19.4
19	Mizoram	65.6	67.0	-	-	-	-	-	34.4	33.0
20	Nagaland	29.4	40.7	-	-	NPF	70.0	54.6	0.6	4.7
21	Orissa	32.5	39.4	16.9	23.1	BJD+	41.7	36.9	8.9	0.6
22	Punjab	45.2	49.8	43.9	45.1	BSP	5.7	3.3	5.2	1.8
23	Rajasthan	47.2	53.9	36.6	40.9	-	-	-	16.2	5.2
24	Sikkim	29.6	16.6	1.8	4.4	SDF	63.3	53.7	5.3	25.3
25	Tamil Nadu	42.5	48.5	2.3	1.5	ADMK+ DMDK	37.4 11.1	32.8 11.5	6.7	5.7
26	Tripura	30.7	28.0	3.4	1.8	CPM	61.7	67.4	4.2	2.8
27	Uttar Pradesh	18.2	19.8	20.8	23.0	SP BSP	23.3 27.4	23.0 29.9	10.3	4.3
28	Uttarakhand	43.3	42.2	34.0	37.7	BSP	15.3	14.5	7.4	5.6
29	West Bengal	44.7	39.4	6.1	5.2	LF	43.4	52.6	5.8	2.8
	All India	36.4	38.8	34.1	26.6	LF BSP SP	7.6 6.2 5.1	6.7 6.6 4.0	10.6	17.3

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:

State electorate proportion weight*(actual turnout figures reported by the ECI ÷ Turnout figures in the state electorate proportion weighted NES 2009).

Vote share weight: As there is a discrepancy in the actual vote share of parties in the Lok Sabha elections and the vote share that we get from the NES 2009 dataset, we correct this by applying a vote share weight. The weight for vote share is calculated by using the following formula:

State electorate proportion weight*(actual vote share for party reported by ECI÷Vote share for the party in the state electorate proportion weighted NES 2009).

Table A below shows both the actual vote share and the state electorate proportion weighted vote share in NES 2009 that was used to calculate the weights for the NES 2009.

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/jati/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
VIII) Others

Table A: Actual Vote Share and Vote Share in NES 2009

	Actual Vote Share	Vote Share in NES 2009
Congress	28.6	31.9
Congress allies	7.8	6.9
BJP	18.8	21
BJP Allies	5.3	5.6
Left	7.6	6.7
BSP	6.2	6.6
Fourth Front	5.2	6.4
Others	20.5	14.8

All figures are percentages. Actual vote share figures are from the ECI. NES 2009 figures are from the state electorate proportion weighted data.

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 example 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<=599)) Scaste=5.
If (Scaste=0 and Z8=1 and (Z7>=600 and Z7<=699)) Scaste=6.
If (Scaste=0 and Z8=2) Scaste=7.
If (Scaste=0) scaste=8.
Format Scaste (F1.0).
Value label Scaste 1 '1: Upper caste' 2 '2: Peasant Proprietors' 3 '3: Upper OBC' 4 '4: Lower OBC' 5 '5: Dalit' 6 '6: ST' 7 '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: 1001 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.

Syntax 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 (z16i>0 or z16j=1) Class=1.
If (Class=0 and (z16f=1 and z16h=1 and z16c>1 and z16k=1 and Lpg=1)) Class=1.
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 z16k=1)) Class=2.
If (Class=0 and (z16c>=1 and z16h=1 and z16k=1)) Class=2.
If (Class=0 and (z16f=1 and z16h=1 and z16k=1)) Class=2.
If (Class=0 and (Income=6 or Income=7)) Class=2.
If (Class=0 and (z16e=1 and z16d=1 and Z16a=1)) Class=3.
If (Class=0 and (z16e=1 and z16d=1 and Z16b=1)) Class=3.
If (Class=0 and (z16d=1 and Z16a=1 and z16b=1)) Class=3.
If (Class=0 and (Income=3 or Income=4 or Income=5)) Class=3.
If (Class=0 and (Z16d=1 and Z16a=1)) Class=4.
If (Class=0 and (Z16d=1 and Z16b=1)) Class=4.
If (Class=0 and (Z16a=1 and Z16b=1)) Class=4.
If (Class=0 and (Income=2)) Class=4.
If (Class=0 and (z16a=0 and z16b=0 and z16c=0 and Z16d=0 and z16e=0 and z16f=0 and z16g=0 and z16h=0 and z16i=0 and z16j=0 and z16k=0)) Class=5.
If (Class=0 and (income=1)) Class=5.
If (Class=0) Class=0.
Format Class (F1.0).
Value label Class 1 'Rich' 2 'Middle' 3 'Lower' 4 'Poor' 5 'Very poor'.
Variable label Class 'Economic class by assets and income'.
execute.

1 This total N is tentative pending the final cleaning of the data that is currently underway.
2 Any dalit or adivasi from a different religion, for example Sikh dalits or Christian adivasis, are placed in the others category according to this coding.
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* (Z7a 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.[7]

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 empirical data and contribute to the theorisation based on these empirical data. Secondly, Lokniti also aims at enriching the practice of political science in India by ensuring that data thus collected is used for writings on Indian politics by researchers and students of Indian politics based in Indian colleges, universities and institutions of research. We hope that this will strengthen social science in India. With this in view, Lokniti has been coordinating the writing and publication of its election analyses since the mid-1990s. *Economic & Political Weekly* has been one of the chief platforms that facilitated these publications (EPW: January, 1996, August 1999, December 2004, February 2009). The collection put together in this issue of EPW is a continuation of this and also a reaffirmation of the goals set out above.

NOTES

1   See Lokniti Team's article entitled "National Election Study 2004: An Introduction" in the special issue of the EPW (18-24 December 2004) for a more comprehensive discussion on these three generations of the NES. The present note in many ways updates this previous article.

2   An election survey was conducted in the union territory of Chandigarh. But, as it was carried out after the elections and as it does not form part of the NES data that is analysed in this issue, it has been included here as a union territory which was not surveyed.

3   The all India rate of substitution in the 2004 NES was 8.5%. However, this varied widely from 0% in Sikkim to about 20% in Goa and Mizoram. See Lokniti Team (2004) for a discussion on this substitution procedure and its variance across states.

4   The achieval rate in 2009 varies from 37% in Arunachal Pradesh to 85% in Karnataka.

5   The questionnaires and related material for the NES 2009 are available at http://www.lokniti. org/national_election_studies.html

6   However, for the PSs that went to the polls in the last two phases of the election, one or two FIs conducted the interviews in each PS to ensure a faster and smoother collection of data.

7   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 draw on this data. For further details visit our website at www.lokniti.org.

REFERENCES

Economic & Political Weekly (1996): 31/2-3, 13-20 January.
  – (1999): 34/34-35, 21-28 August.
  – (2004): 39/51, 18 December.
  – (2009): 44/6, 7 February.
Jaffrelot, Christophe (1996): *The Hindu Nationalist Movement and Indian Politics* (New Delhi: Penguin).
Journal of Indian School of Political Economy (2003): 15/1-2, January-June.
Kumar, S, A Heath and O Heath (2002a): "Determinants of Social Mobility in India", *Economic & Political Weekly,* 37: 2983-87.
  – (2002b): "Changing Patterns of Social Mobility: Some Trends Overtime", *Economic & Political Weekly,* Vol 37, No 40: 4091-96.
Lokniti Team (2004): "National Election Study 2004: An Introduction", *Economic & Political Weekly,* 18 December, 5373-82.
Mitra, S K and V B Singh (1999): *Democracy and Social Change in India: A Cross-sectional Analysis of the National Electorate* (New Delhi: Sage).
Shastri, Sandeep, K C Suri and Yogendra Yadav (ed.), (2009): *Electoral Politics in Indian States* (New Delhi: OUP).
Vaid, D and A F Heath (forthcoming): "Unequal Opportunities: Class, Caste and Social Mobility" in A F Heath and R Jeffery (ed.), *Diversity and Change in Modern India: Economic, Social and Political Approaches* (Oxford: Oxford University Press), Proceedings of the British Academy, 159.
Varshney, A (2003): *Ethnic Conflict and Civic Life: Hindus and Muslims in India* (New Haven: Yale University Press).
Yadav, Yogendra (1999): "India's Third Electoral System", *Economic & Political Weekly,* 34/34-35, August, pp 2393-99.
  – (2004): "The Elusive Mandate of 2004", *Economic & Political Weekly,* 39/51, 18 December, pp 5383-98.
Yadav Yogendra and Suhas Palshikar (2003): "From Hegemony to Convergence: Party System and Electoral Competition in Indian States, 1952-2002", *Journal of Indian School of Political Economy,* 15/1-2, January-June, pp 5-44.
  – (2009): "Principal State Level Contests and Derivative National Choices: Electoral Trends in 2004-09", *Economic & Political Weekly,* Vol 44, No 6, 7 February, pp 55-62.

Appendix 1: Lokniti Research Team

Sanjay Kumar of CSDS, Delhi coordinated this All India Survey. The research team at Lokniti that coordinated and analysed the survey comprised Sanjeer Alam, Banasmita Bora, Vanita Leah Falcao, Vikas Gautam, Dhananjai Joshi, Navprieet Kaur, Praveen Rai, Ram Narayan Raman, 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

(1) Assamese, (2) Bengali, (3) English, (4) Garo, (5) Gujarati, (6) Hindi, (7) Kannada, (8) Khasi, (9) Konkani, (10) Malayalam, (11) Manipuri, (12) Marathi, (13) Mizo, (14) Nepali, (15) Oriya, (16) Punjabi, (17) Tamil, (18) Telugu, (19) Tripuri, and (20) Urdu.

Appendix 4: List of State Coordinators

State	Name of the Coordinator
Andhra 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 Jani 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	Sajad Ibrahim
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 Mishra
Punjab	Ashutosh Kumar and Jagroop Singh Sekhon
Rajasthan	Sanjay Lodha
Sikkim	Mukund Giri
Tamil Nadu	Gundapuneni Koteswara Prasad
Tripura	Sukhendu Debbarma
Uttar Pradesh	Mirza Asmer Beg, Sudhir Kumar and A K Verma
Uttarakhand	Annpurna Nautiyal
West Bengal	Suprio Basu

Appendix 5: Data Unit Team

Himanshu Bhattacharya, K A Q A Hilal and Kanchan Malhotra of the CSDS Data Unit.