Psephological Fallacies of Public Opinion Polling

PRAVEEN RAI

Opinion polls in India capture electoral snapshots in time that divulge information on political participation, ideological orientation of voters and belief in core democratic values. The survey data provides for crucial social science insights, validation of theoretical research and academic knowledge production. Although the media's obsession with political forecasting has shifted to electoral prophecy, psephology continues to provide the best telescopic view of elections based on the feedback of citizens. The ascertainment of subaltern opinion by surveys not only broadens the contours of understanding electoral democracy, but also provides an empirical alternative to the elitist viewpoint of competitive politics in India.

Praveen Rai (*praveenrai@csds.in*) is a political analyst at the Centre for the Study of Developing Societies, Delhi.

The terms "survey" and "opinion poll" in India would have remained a professional jargon of market research industry, had it not been used for predicting election outcomes. The green shoots of opinion polls to study Indian national elections emerged in the 1950s, but it caught the imagination of the people and became clichéd in the closing decade of the 20th century. The popularity of election surveys stems from the political socialisation and crystal ball gazing curiosity of Indians to foresee the outcomes of hustings before the pronouncement of formal results. The electoral inquisitiveness of the stakeholders created a large canvas of opportunity for opinion-polling industry and scope for scientific forecasting of Indian election competitions. The proliferation of electronic media and the rapid monetisation in the 1990s provided momentum to polling agencies to venture into opinion polling on national electoral politics and state election contests. The opinion polls captured panoramic snapshots and divulged the sociodemographic characteristics of Indian voters and their nuanced voting preferences, as well as reasonably accurate vote estimates of political parties for predicting elections. The fixation for survey-based election prediction turned a host of political scientists and television anchors in India into psephologists. The media election soothsaying became so definitive and encapsulating that it overshadowed the announcements of election results by the Election Commission of India (ECI) to a mere formality of medal distribution ceremony. The success of mediatised election prophecy was short-lived, however, as the erroneous prediction of the national election results in 2004 led to widespread public criticisms and calls for a blanket ban of pre-poll, exit poll and post-poll election surveys during elections in India.[1]

The media-opinion polling industry, facing an existential threat, resorted to course correction, but the election polling ecosystem turned from bad to worse between 2005 and 2013, as political parties were caught manipulating in-house election survey data for mobilisation of the electorate. This marked a tectonic shift in purpose, as initially political outfits commissioned election surveys to gauge the mood of voters, collect grassroots feedback for selection of "winnable" candidates and formulation of election strategy and manifesto. Between 2014 and 2019, the correct election forecast of state elections by opinion polls led to rebuilding the confidence of people in quantitative analysis of election and balloting. The polling industry, with increased accuracy in election forecasting, faced a crucial litmus test in general elections 2014, as the incumbent Congress-led United Progressive Alliance (UPA) was up against a resurgent and combative Bharatiya Janata Party (BJP)–led National Democratic Alliance (NDA). The election prediction of

most market research agencies correctly fathomed the direction of the 2014 mandate, but none of them (except one poll) could forecast a BJP majority of seats for the first time in the lower house of Parliament. The Lok Sabha elections in 2019 once again posed a major challenge for polling agencies, as most Indian and foreign media could not find the prevalence of the "Modi wave" and predicted with aplomb that it will be a waveless national husting. The exit poll findings were diametrically opposite of the media narrative of the fading saffron wave, as it revealed that the BJP-led right wing alliance would do an impressive electoral rebound with a bigger mandate. The political forecast by pollsters proved correct, but barring two polling agencies, none could prophesise that the BJP would win more than 300 out of the 543 Lok Sabha seats. The prediction of the people's mandate by election surveys was in the right direction, but most of them once again failed in correct assessment of the magnitude of the BJP's political triumph.

Thus, it becomes important to revisit and review the opinion polls for quantitative analysis of election and balloting from a holistic perspective to fathom their intrinsic limitations, situational challenges and statistical accuracy errors in predicting election outcomes. This article comprises of three major parts. The first collates the accuracy levels of public opinion polls by media conglomerates of national elections between 1998 and 2019 in India. A comparative analysis of seat prediction based on election surveys with national election verdicts will not only help in foregrounding the precision levels achieved by the polling industry in the last two decades, but also map the degree of standard deviations. The archiving will provide the growth continuum of election opinion polls and decipher the changes in psephological trajectory and the methods for summating public opinion. The second part theoretically fathoms the biases and fallacies of psephology and the insurmountable challenges faced by the polling organisations in conducting accurate election surveys. It discusses the inherent errors and bias in survey research and their impact on the accuracy quotient of election studies in the multiparty structure of Indian competitive politics. The final part deconstructs the fundamental flaws in opinion polls, scrutinises the reasons for erroneous poll predictions and probes the media sensibility in distinguishing between empirical and anecdotal evidence while reporting election data analytics. It will delineate the psephological advancement and traction of public opinion polls along with transmutation of purpose from understanding elections to a media utility for political predictions.

Accuracy of National Election Opinion Polls

Poll accuracy in the initial years of opinion polling was a measure of closeness of vote estimates of political parties (observation) with the vote share (true value) figures of the ECI. However, due to paramountcy of election forecasting in India, the paradigm of survey accuracy witnessed a shift from comparison of vote shares to closeness of seat predictions regarding the number of seats won by political parties. Predicting elections became an integral part of public opinion polling and established primacy over mapping of electoral behaviour and attitudes. A tension always existed between survey researchers who produced pre-election opinion polls and consumers who watched or read poll stories during election campaigns. The pollsters emphasised that their results are only an election "snapshot in time" and not a "political forecast." While some pollsters dismiss the notion of treating pre-election polls primarily as a forecasting instrument, many in the field are happy to treat the apparent accuracy of polling near elections as a sign of its overall health. It became implicit in the scoring of poll accuracy, in Crespi's (1988) resolution long ago, that ruling out the forecasting value conducted "immediately before an election" is to "impugn meaningfulness of all polls." If polls cannot achieve such predictability, why should we accept any poll results as meaning relevant to real life (Crespi 1988; Blumenthal 2014)? It thus becomes contextual to compile and engage in a comparative analysis of poll predictions with the national elections results to ascertain the accuracy rates of public opinion polling in India.

The general elections in 1998 marked the mushrooming of opinion polls to read the minds of voters and the beginning of a media rat race for seat predictions and eureka moments. The majority of election polls predicted the seat tally of the BJP allies quite closely, except *India Today*–CSDS (Centre for the Study of Developing Societies), which underestimated the victory margin of the saffron party and its political partners by a wide margin, but surprisingly got it right for the losing Congress alliance. The explanation for some discrepancies between the predicted and actual seats can perhaps be on the gap between the date of the election survey and the dates of the actual voting. The design of the opinion poll was primarily to ascertain an all-India prediction of seats, and though the CSDS method involved predicting seats for all states, the sample size at the state level was too small to offer any measure of confidence in the predictions (Karandikar et al 2002).

The next big test for election forecasting was the general elections in 1999, which again proved to be a red-letter day for the polling industry. The polls correctly predicted a victory of the BJP alliance, with a slight overestimation of seats for the winning political combine. The election predictions based on surveys done by various media houses during the general elections in 1998 and 1999 were quite accurate and almost mirrored the election results. It received public kudos and appreciation and boosted the growth of opinion polling industry in India (Rai 2014).

Table 1: General Elections 1998: Almost Accurate Prediction

Seat Forecast	BJP Allies	Congress Allies	Others
DRS-*Times of India**	249	155	139
Outlook-C Nielsen	238	149	156
India Today-CSDS	214	164	165
Frontline-CMS**	225–235	145–155	152–182
Actual result	252	166	119

Seat predictions based on opinion poll conducted before the elections.

Table 2: General Elections 1999: NDA Tally Slightly Overestimated

Seat Forecast	BJP Allies	Congress Allies	Others
Times poll-DRS	332	138	-
Outlook-CMS	319–329	135–145	34–39
India Today-Insight	332–336	132–146	70–80
HT-AC Nielsen	300	146	95
Pioneer-RDI	313–318	140–150	84–86
Actual result	296	134	113

Seat predictions based on opinion poll conducted before the elections.

The near accurate political prediction by the election survey

industry in the last decade of the 20th century provided an opportune time to improve its scientific validity and election forecasting models. The 2004 general election in India was a big moment for opinion polling to stamp its reliability, but proved to be its Achilles heel, as the election predictions turned out to be completely upside down. The media polls predicted, with an air of self-confidence, that the incumbent BJP-led NDA alliance will retain power and be at the helm of affairs for the next five years. The various polls differed on the number of parliamentary seats the saffron combine would win, with some suggesting that it would return with a bigger tally, while others predicted some losses. The 2004 hustings proved to be a topsy-turvy election, as the NDA alliance led by one of India's most popular Prime Ministers, Atal Bihari Vajpayee, was defeated by the Congress-led alliance in a shocking turn of political events.

A dissection of seat predictions in the 2004 general election, disaggregated at state level, reveals that the psephological shocks in Andhra Pradesh and Tamil Nadu had a multiplier effect on intensifying survey bias and margin of error leading to an erroneous election forecast. The structural reasons for election prediction missing the bull's eye is mainly due to the nature of Indian polity that includes a first-past-the-post (FPTP) electoral system, multiparty election competition and high volatility of the electorate. The challenge of converting vote percentage into seats that confronts pollsters in India and why this flawed representative feature, in a built-in FPTP system, accounts for some measure of inaccuracies in a number of opinion and exit polls (Butler et al 1995). The functional limitations of election survey-based forecast is due to the ideological bias of the media, statistical rebalancing by pollsters, improper political interference, data tweaking for client needs, sample size inadequacy and unscientific survey methodology. The issue of methodology, rigour and underestimation bias (play-safe mode) on the part of pollsters is due to the need to predict the winner correctly, as getting it wrong incurs dangers of disrepute. The pollsters "play-safe" to get the winner right, and prefer to err on the seats by using the lower bound of seats as predicted by their statistical models. Interestingly, along with the pollsters' play-safe approach, the exact seat prediction becomes difficult also because a sizeable number of respondent voters also prefer the play-safe approach in their replies to survey questions on their choice of political party (Butler et al 1995). The stakes are so high for data aggregators during the elections that they stick their neck out only for predicting the winning political party/parties and now refrain from stating the exact scale of victory.

The polling industry, facing public heat and a clarion call for a blanket ban in election survey after the debacle of the 2004 elections, got another opportunity in 2009 to ward off allegations of influencing votes and vitiating the ecosystem of free and fair elections. The opinion polls predicted a tough fight between the BJP and Congress-led political alliances and a hung assembly with a slight edge for the Congress-led combine. It completely misread the upcoming verdict and failed to fathom the Congress party upsurge and substantial gains in number of parliamentary seats from the previous election. Five years down the line, the poll predictions by different media houses and pundits in 2009 once again failed to predict the victory of the incumbent Congress-led UPA government. The only poll that predicted that the Congress party would win around 200 Lok Sabha seats was its own internal survey, but it did not meet public approval. The negative questions that arose after the 2004 Lok Sabha elections resurfaced, and the credibility ratings of the opinion-polling industry severely nosedived. The average record of pre-poll opinion surveys in 2004 and 2009 Lok Sabha elections reveals the implicit weaknesses and dangers of data rigging in election forecasting in a neck-to-neck electoral competition. In such a complex political setting, leading public misconceptions ranged from election survey being an unscientific exercise, to an instrument of political communication covertly used by media houses for false predictions, aimed at influencing the non-opinionated electorate in India (Rai 2014).

The reasons for inaccurate seat predictions of the 2009 Indian general elections for the two dominant political party combinations in the fray are identical to those of the 2004 national elections debacle. The distortion between vote shares and election results in terms of seats—due to the voting system—is not a new phenomenon in India, as the electoral fate of parties always depends more on their adversaries' electoral performances than on their own. The multiplication of triangular-and sometimes quadrangular-competitions rendered any attempt at predicting the electoral outcome in the 2009 national elections completely illusory, due to increased fragmentation of the political scene. This optical illusion stems from the single-round majoritarian voting system, which meant that the growing fragmentation of the regional political scenes, particularly in the case of triangular or quadrangular competitions, acted in favour of the Congress, the most consistent, though not dominant player (Jaffrelot and Gilles 2009).

Table 3: General Elections 2004—UPA Seats Grossly Underestimated

Seat Forecast	BJP Allies	Congress Allies	Others
NDTV–*Indian Express*	230–250	190–205	100–120
Aaj Tak–ORG Marg	248	190	105
Zee-Taleem	249	176	117
Star–C–Voter	263–275	174–186	86–98
Sahara–DRS	263–278	171–181	92–102
Outlook–MDRA	280–29	159–169	89–99
Actual result	189	222	132

Seat predictions based on opinion poll/exit poll conducted during the elections.

Table 4: General Elections 2009: Failed to Forecast Congress Upsurge

Seat Forecast	BJP Allies	Congress Allies	Others
STAR News–AC Nielson	197	199	136
CNN-IBN	165–185	185–205	165–195
NDTV	177	216	150
Headlines Today	180	191	172
News X	199	191	152
Times Now	183	198	162
Actual result	159	262	79

Seat predictions based on opinion poll/exit poll conducted during the elections.

The political churnings seemed quite visible during the 2014 Lok Sabha elections. The political arena had changed diametrically for the Congress-led UPA-II government and it faced twin electoral disadvantages: a strong anti-incumbency, and an amplified "[Narendra] Modi wave." The opinion polls predicted that the NDA led by the BJP would show the UPA the exit doors and come back to power in

Delhi, which came true. The correct prediction of the people's mandate brought a sigh of relief for pollsters and their part failure was lost in the celebratory din. A glance at Table 5 shows that (except one polling agency) none of the opinion polls could predict that the BJP, riding high on a saffron wave, could cross the majority mark on its own—272 out of 543 parliamentary seats—in the lower house. The correct assessment of the national election verdict in 2014 reinvigorated the survey aggregation industry and provided traction to carry out improvements in computational capabilities. The scientific and technological advances made by the opinion polling industry resulted in greater public visibility and higher accuracy of poll predictions in state elections between 2014 and 2018.

The 2019 Lok Sabha election was one of the sharpest ideological electoral competitions between the right-wing BJP alliance and the centre-left as well as left-of-centre party combine with incompatible political visions. The election writings by prominent political scientists and reporting by seasoned journalists indicated that the Modi wave was absent and it would certainly be a "wave-less" national election. They acknowledged that the saffron party alliance had a slight electoral edge over other political parties, which may or may not be enough to catapult the BJP and its allies back to power in Delhi. Academic and media experts found solace in pre-election surveys that played safe, and forecast that the BJP and its political partners were losing electoral steam and may fall short of the majority mark needed to form the government. The exit polls after the last phase of the election and before the results revealed that BJP alliance would cross the 300-seat mark, with two polling agencies hitting the jackpot by predicting 350-plus seats. The correct predictions reinstated opinion polling in India to its pride of place as the most authentic source of election information and political prediction. The psephological success not only vindicated the accuracy of empirically evidenced election assessments, but also trashed the political writings of public intellectuals and fourth estate specialists that predicted a doomsday election for the Modi-led BJP and its tireless quest for a re-election.

Table 5: General Elections 2014: Failed in Predicting a Majority for BJP

Seat Forecast	BJP Allies	Congress Allies	Others
ABP–AC Nielson	281	97	165
CNN-IBN–CSDS	276–282	92–102	150–159
Headlines Today-CICERO	261–283	101–120	152–162
India TV–C-Voter	289	101	153
News 24–Today's Chanakya	340	70	133
Times Now–ORG	249	148	146
Actual result	326	60	157

Seat predictions based on opinion poll/exit poll conducted during the elections.

Table 6: General Elections 2019: Correct Estimation of NDA 3.0 Seats

Seat Forecast	BJP Allies	Congress Allies	Others
India Today-AXIS My India	339–365	77–108	69–95
Today's Chanakya	350	95	97
News18-Ipsos	336	82	124
Times Now-VMR	306	132	104
India News	298	118	127
Republic-CVoter	287	128	127
ABP Nielsen	277	130	135
Actual result	353	91	98

Seat predictions based on opinion poll/exit poll conducted during the elections.

Thus, the analytical summary of survey-based election forecasting of the national elections reveals a mixed bag, as it started with bang on predictions in 1990s, floundered in the first decade of the 21st century and regained its lost ground in the last two Lok Sabha elections. The accuracy of election forecasting at the national and state elections in India has considerably improved in recent times; unfortunately, incorrect political predictions stick in the public memory, creating a perception deficit about their reliability, competence and neutrality (Kumar et al 2016). Thus, it becomes pertinent to dig deeper into election survey discourse to find the inherent fallacies and statistical limitations of polls and its adverse impact on elections forecasting.

Fallacies of Election Opinion Polls

The study of elections is perhaps more challenging in India than other democratic countries, as it involves understanding the interplay of overlapping social cleavages in one of the world's most demographically heterogeneous countries. The high political fluidity in a multiparty system and the complexities of a developing economy pose serious difficulties in the correct forecasting of elections. Moreover, contrary to what many believe, opinion poll-based studies of elections are not just about picking likely winners and computing margins of victory/defeat, but also provide a nuanced understanding of the different factors that go into determining the election verdict. It provides quantitative evidence on why Indian people voted the way they did, the changes and continuities from the past, and what the future might look like for its citizens (Kumar et al 2016). The FPTP voting system and multiparty electoral competitions makes it quite daunting for opinion polls to gather precise vote shares as compared with countries that witness bipolar election contests. In multipolar contests with three or more dominant political parties, a slight error in vote approximation can completely upset the apple cart of seat predictions. Similarly, the parties that contest elections in political partnership or through formation of new alliances also pose a serious problem, as election surveys cannot capture the working of the alliance at grassroots and transfer of committed votes to each other. The uneven concentrations of votes for some parties in some regions and spatial pockets also make it difficult to obtain the right election forecast, even if the vote share estimates of the main political dispensations are correct. The election opinion polls are quite limited in focus as they cannot measure the merger and split of political parties, political heavyweights changing affiliations, factionalism in parties, influence of rebel candidates and the localised dynamics of electioneering.

The profiling of the Indian electorate reveals that it is highly heterogeneous and complex. The multiple identities of voters on regional, caste community, linguistic and religious identity overlap and make it difficult to ascertain the patterns and continuity of their political affiliation. The Muslims in India are an apt illustration of heterogeneity marked by regional, linguistic, sectarian and spatial differentiations, reflected in their political

choices and explaining the absence of en bloc vote for a particular political party, which debunks the vote bank theory of Indian electoral politics. Similarly, the voting behaviour of the electorate has been quite volatile, as they switched political allegiance between two elections on several occasions. The CSDS election studies data reveals that around one-fourth of the voters finalise their voting choice after the election campaign is over (floating voters). The pre-election opinion poll forecasts generally go wrong, as they cannot factor in the voting intention of the floating voters. The intrusive surveillance of the electorate in India through personal visits by political party cadres and social media platforms in recent times has also created a "fear of reprisal" electoral environment. Hence, voters with threat perception of fear desist in spelling out the correct voting preferences and play safe by naming the dominant party in their locality. There is a lurking fear that survey research on political opinions and attitudes can serve as an instrument—more effective and therefore more dangerous—of neocolonial surveillance of global populations (Yadav 2008).

The most critical functional component of election surveys is the selection of survey design and the sampling methodology as it not only determines the quality and accuracy of the survey but is also a precursor for foretelling the likely election outcomes. The samples drawn for CSDS polls are from the voter lists available with the ECI and use multi-stratified probability-sampling technique. The sample selection is generally accurate and yields a representative sample avoiding both the coverage and sampling error. On the other hand, media opinion polls rely on quota sampling for estimating vote share and election prediction. It involves a fixed number of respondents based on gender, education, caste communities and age differentials. This results in a skewed profile of voters—with coverage and sampling errors in abundance—with incorrect vote shares. The polls suffer from urban sample bias due to high costs and logistical inconvenience in reaching far-off villages and remote habitations. If a sample survey fails to gather the opinion of any important caste and community, the election predictions will be highly vulnerable to failure. The method of sample selection is crucial for election surveys and most Indian polls go wrong because their sampling methodology is poor, which makes the sample profile unrepresentative. Though a scientific and representative sample determines the accuracy of the survey, there is no guarantee that a forecast based on the survey will be right. A survey has its limitations, as it cannot capture the diverse and nuanced complexities and undercurrents of electoral behaviour and choices in India (Yadav 2008).

The five basic procedures of opinion polling usually carried out in the ascending sequence are as follows: First, questions are written and organised into questionnaires; second, a sample is selected to represent the population to be surveyed; third, designated respondents are interviewed; fourth, answers given are statistically analysed; and, fifth, results are interpreted and conclusions reached (Young 1990). The errors that creep in opinion polls mostly arise in the first two stages, that is, framing the design of the interview schedule (questionnaire) and drawing a scientific sample from the universe of study. The fallacies that arise from question wording in opinion poll questionnaires include inappropriate and concealed use of persuasive definitions, broad definitions to inflate statistical results, meaningless statistics, atypical examples in place of a definition, question structure bias, dichotomous questions, and double negatives in question wording. The fallacies are especially deceptive in polls because of the misleading appearance of objectivity encouraged by the ways polls are typically presented to the consumers. The persuasive spin on the question is concealment by the objective appearance of the announced poll result, especially when it is presented with a numerical calculation of the probability of error. These fallacies point to a deeper problem in evaluating bias in the questioning used in polls. The fallacy of question structure bias in polling is committed where the structure of the question introduces a bias into the poll that is deceptive, concealing a "yea-say effect."[2] This tactic is often employable in push polling in political polls, as the question in use is really an attempt to persuade or to influence voters, but on the surface, it appears as a routine parade for merely collecting information by taking a poll (Walton 2007). The market opinion polls need to address and overcome the structural and functional constraints and ascertain the right vote share estimates, but it depends upon the projection model to convert it into the correct number of seats that political parties are likely to win.

Fallibility of Political Predictions

As noted, the forecasting of elections in a country like India is complicated and difficult due to a populace comprising of myriad caste community groupings, combined with multiple political parties across the political spectrum. An election survey can estimate the vote shares correctly for the political parties, but predictions can still go wrong due to intrinsic flaws in forecasting models or due to pollsters tweaking projections based on statistical wisdom or rebalancing by media to suit their political preferences. The prediction of election results is a relatively recent and increasingly popular part of political science research. Competitive elections are the hallmark of modern democracy and being able to foreshadow who wins them is a tantalising skill that has garnered significant scientific attention (Jackman 2005). Election forecasting stands out from many other kinds of political science research in a number of ways. It is highly data-driven, focused on a very concrete and delimited task, and in most studies, the goal is not to explain election outcomes but to describe and predict them. In that sense, the question of "how" rather than the standard scientific question "why" is in focus. The question "how" is still highly relevant from a scientific perspective, and to achieve reasonable accuracy the need is to make the most out of the limited and flawed polling data, controlling seasonal fluctuations in public opinion, variability in measurements and bias associated with polling houses (Walther 2015).

The parliamentary/state assembly seat projection models used by pollsters in India for forecasting elections are primarily based on a statistical method, the "probabilistic count" that

uses swing of votes between two elections. The socio-economic factors influence the voting patterns significantly, but it is reasonable to assume that the socio-economic profile of most of the constituencies does not change significantly from one election to the next. Therefore, while the differences in socio-economic profiles between the two constituencies are a reflection in voting patterns in a given election, the change from one election to the next in a given constituency does not depend on them. The change in the percentage of votes for a political party in an Indian state from the previous election to the present is assumed to be constant. The change in the percentage of votes is the "swing" factor. Under this model, the big states are divided into geographic regions with the postulation that the swing in a seat is a convex combination of swing across the state and swing across the region (Karandikar 2014). The model uses the votes share information for the previous election from the ECI archives, gathers the vote estimates for the current election by a sample survey, and applies an even swing of votes for all the political parties in the fray across all the state assembly constituencies. The uniform swing of votes seldom happens, as only some constituencies may witness bigger shifts. The model also fails to take care of the uneven vote concentration in some regions and the marginal victories of political parties in a number of seats. The model is not very accurate if we look at historical data, but is a reasonably good approximation to predict the seats for major political parties at national level elections (Karandikar 2014).

A sample of 4,000 voters in 543 parliamentary constituencies can predict the seats accurately, but a sample size of over 21 lakh would be impractical, as it will entail a huge cost and require an army of trained and reliable field enumerators. Thus, predicting seats on a mathematical model of vote shares ascertained at state level from a cluster of 10–12 assembly constituencies increases the possibility of modelling error. The other limitation of a survey done well ahead of the actual polling day is that though it measures the opinion of the whole population, what really counts is the group that actually goes out and votes. The CSDS election data reveals that the propensity to vote is much lower among the urban, upper middle class and upper class, college-educated and high-income groups. The electorate is quite volatile and voting intentions undergo massive swings as voting day approaches in India. These two factors mean that the predictive power of any election opinion poll done weeks in advance is limited and fallible, as all it can measure is the mood of the nation at the time of the poll (Karandikar 2014). Voters may change their minds between an opinion poll and the election day, and this is the main reason why polls taken six months before an election have a much poorer predictive record than those taken close to the election date (Northcott 2015). The traditional polls are snapshots of public opinion at a certain point in time and do not provide predictions. The routine interpretation of polling results as election day forecasts can result in poor predictions, particularly if the election is still far way, because public opinion can be difficult to measure and remains fragile over the course of an election campaign (Campbell 1996).

The voting intentions of a sample serve as a proxy for those of a population and the main reason for an unrepresentative sample is the sampling error, as small samples can lead to misleading flukes. A major issue for pollsters is to ensure that the samples are in appropriate balance with respect to various demographic variables, and if required, use balancing procedures to put relevant weights. In addition to sampling errors and systematic bias, the phenomenon of "herding" can lead to forecasting error. It is widely suspected that most polling agencies, at the end of a campaign, "herd" and report headline figures closer to the industry mean, presumably to avoid the risk of standing out as having missed the final result by an unusually large margin. Some sensitivity to this turns out to be optimal for accurate election prediction (Northcott 2015). The vote share estimates of an election poll and seat predictions can be fully correct, but as part of media manipulation, different figures could be publicly released and the subsequent error be blamed on a faulty projection model. The main problems concern the unwarranted and misleading inferences drawn from polls by their readers and users—often an audience that may not be well aware of the limitations of statistical methodology. There are several statistical polls run by the media under the pressure of deadlines and to puff up a poll by published findings that may excite readers. There is no surety that opinions polls are fallacy-free and the onus is on a critical thinking public to become aware of the biases and fallacies, and to assume a "buyer beware" attitude. Social statistics are needed to conduct intelligent public deliberations and set social policies in a democracy, but activists, the media and private agencies can and often do use "mutant statistics" as tactics to manipulate public opinion (Best 2001).

In the absence of forecasting models of Indian polling firms in the public domain, it is quite difficult to assess the status of research and development, but the increased accuracy rates of political predictions in the last decade indicates its statistical advancements. In contrast, the United States (US), from the 1970s onwards, witnessed an addition of a wide range of successful election forecasting techniques in the literature on electoral forecasting. It is clear that psephological improvements in the models of fundamentals is an unpromising route and might involve getting better polling data, analysing that data better, or understanding better how the implications of that data depend on local peculiarities and grassroots politics.

Conclusions

To conclude, stocktaking of opinion polling in the last 40 years reveals that 75% of the 833 (386 pre-poll and 447 exit) election surveys correctly predicted the winning political party (or parties) in India. The accuracy rate of exit polls (84%) was 13 points higher than opinion polls (71%) conducted during the elections. The success rates—aggregate of both exit and opinion polls—of polls differ quite significantly for the national and state elections. The correct prediction for Lok Sabha elections is 97%—the 2004 Lok Sabha polls was an outlier—while the success rate is 75% for state assembly polls. The strike rate of such polls may not match the global standards of the polling industry, but they are not as off the mark as public perception

imagines, hence it is perilous to dismiss opinion polls (Roy and Sopariwala 2019). The election forecast record of Indian polling firms may not match the world benchmark, but a post-mortem of election polling reveals that precision in terms of vote share accuracy is at par with its US and United Kingdom contemporaries. The mathematical prediction models based on opinion poll vote share is fallible as well as fragile, but polling agencies in India guided by blue-sky thinking are trying their best to improve the craft of political forecasting and seat predictions.

A US study reveals that vote expectation surveys provide more accurate forecasts of election winners and vote shares than four established methods of studying elections: vote intention polls, prediction markets, quantitative models and expert judgment. The opinion poll-based vote expectation survey is inexpensive and easy to conduct; results are easy to understand and provide accurate and stable forecasts and thus make it difficult to frame elections as horse races. The use of judgment of political insiders and experienced election observers to forecast elections had been in practice long before the emergence of scientific polling, and it is still invaluable. The common assumption is that political experts and media specialists have enormous experience in reading and interpreting polls, assessing their significance during campaigns and estimating the effects of recent or expected events on the aggregate vote. However, given their omnipresence, surprisingly little is known about the relative accuracy of experts' election forecasts (Graefe 2014). The wide off-the-mark prediction of the national elections in 2019 by political scientists, public intellectuals and media experts in India due to their political bias or ivory tower approaches, vindicates opinion polls as the best and most reliable source for study and comprehension of Indian elections.

The utility of opinion polls extends beyond seat and vote share projections, more specifically, to the production of public knowledge. The data from polls help in providing crucial social science insights and have great academic value. Social scientists have used survey research to answer many important questions about the polity and society. The time-series survey data in particular has been useful in studying long-term trends of Indian politics: political participation, ideological orientation of voters, trust in institutions, the efficacy of the vote, degree of belief in a democratic system and leadership choices, to list a few in the field of psephological advancement. The scientific models for forecasting are a common practice in pure sciences and social sciences like economics and sociology. Thus, election survey evidence can also be of use in preparation of similar projection models to understand political and social events like electoral competitions. The empirical models on studying human behaviour are more prone to error, but the intrinsic bias and trust deficit against election forecasting must be curbed, since elections remain the most opportune moment to study politics and people in India (Kumar et al 2016). The obsession of media opinion polls in forecasting elections has shifted the focus from psephology to electoral prophecy and a few wrong seat predictions undermined its value, but it continues to provide the best telescopic view of electoral politics based on opinion and attitudes of common people. The ascertainment of subaltern opinion by election surveys not only broadened the contours of understanding electoral democracy, but also provided an empirical alternative to the elitist viewpoint of competitive politics in India.

NOTES

1 Pre-poll survey, as the name indicates, is an election survey that happens much before balloting to measure popular choices about political parties, contesting candidates and political leaders. It gauges the voting behaviour and attitudes of sampled voters. Exit poll, as its name overtly suggests, is a survey of sampled electorate as they come out or exit from the polling stations after casting their votes. The survey takes place on the polling day, hence it is also known as election day polling. Post-poll survey is an indigenous method of survey for measuring voting behaviour and attitudes in India, pioneered by the Centre for the Study of Developing Societies (CSDS), Delhi in the 1960s. The voters are interviewed after they have exercised their franchise in the relaxed confines of their homes or workplaces. The post-poll opinion poll is purely an academic exercise for doing a quantitative post-mortem analysis of elections (Kumar and Rai 2013).

2 A "yea-say effect" occurs when a question used in a poll is posed in one direction only. The more a question involves a subject "on which knowledge is hazy or about which people have not thought widely," the wider is the margin of agreement over disagreement (Roper 1990).

REFERENCES

Best, Joel (2001): *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*, Berkeley: University of California Press.

Blumenthal, Mark (2014): "Polls, Forecasts, and Aggregators," *Political Science and Politics*, Vol 47, No 2, pp 297–300.

Butler, David, Ashok Lahiri and Prannoy Roy (1995): *India Decides. Elections 1952–95*, Delhi: Books & Things.

Campbell, James E (1996): "Polls and Votes the Trial-Heat Presidential Election Forecasting and Political Campaigns," *American Politics Research*, Vol 24, pp 408–33.

Crespi, Irving (1988): *Pre-election Polling: Sources of Accuracy and Error*, New York: Russell Sage Foundation.

Franklin, C F (2008): "How Pollsters Affect Poll Results," *Political Arithmetik*, 24 August, http://politicalarithmetik.blogspot.com/2008/08/how-pollsters-affect-poll-results.html.

Graefe, Andreas (2014): "Accuracy of Vote Expectation Surveys in Forecasting Elections," *Public Opinion Quarterly*, Vol 78, No S1, pp 204–32.

Jackman, Simon (2005): "Pooling the Polls over an Election Campaign," *Australian Journal of Political Science*, Vol 40, No 4, pp 499–517.

Jaffrelot, Christophe and Verniers, Gilles (2009): "India's 2009 Elections: The Resilience of Regionalism and Ethnicity," *South Asia Multidisciplinary Academic Journal* (Online), Vol 3, https://doi.org/10.4000/samaj.2787.

Karandikar, Rajeeva L (2014): "Elections 2014, Power and Limitations of Opinion Polls: My Experiences," Hindu Centre for Politics and Public Policy, 2 April, https://www.thehindu-centre.com/verdict/commentary/article5739722.ece.

Karandikar, Rajeeva L, Clive Payne and Yogendra Yadav (2002): "Predicting the 1998 Indian Parliamentary Election," *Electoral Studies*, Vol 21, No 1, pp 69–89.

Kumar, Sanjay, Praveen Rai and Pranav Gupta (2016): "Do Surveys Influence Results?" *Seminar,* Vol 684.

Kumar, Sanjay and Praveen Rai (2013): *Measuring Voting Behaviour in India,* Delhi: Sage Publications.

Northcott, Robert (2015): "Opinion Polling and Election Predictions," *Philosophy of Science*, Vol 82, No 5, pp 1260–71.

Rai, Praveen (2014): "Fallibility of Opinion Polls in India," *Economic & Political Weekly*, Vol 49, No 18, pp 13–17.

Roper, Burns W (1990): "Are Polls Accurate?" *The Classics of Polling*, Michael L Young (ed), Metuchen, New Jersey: Scarecrow Press.

Roy, Prannoy and Dorab R Sopariwala (2019): *The Verdict: Decoding India's Elections*, Penguin Random House India.

Walther, Daniel (2015): "Picking the Winner(s): Forecasting Elections in Multiparty Systems," *Electoral Studies*, Vol 40, pp 1–13.

Walton, Douglas (2007): *Media Argumentation: Dialectic, Persuasion and Rhetoric*, Cambridge University Press, pp 228–74.

Yadav, Yogendra (2008): "Whither Survey Research? Reflections on the State of Survey Research on Politics in Most of the World," Malcolm Adiseshiah Memorial Lecture, Chennai.

Young, Michael L (ed) (1990): "Editor's Introduction to the Basics of Polling," *The Classics of Polling*, Metuchen, NJ: Scarecrow Press, pp 47–50.