Learning From Exit Polls In Sequential Elections*

Manasa Patnam[†]

CREST (ENSAE)

Abstract

This paper examines social learning in voting behaviour through the release of exit poll forecasts amid sequential multi-phased elections. Using both aggregate and post-poll survey data on voting behaviour together with poll predictions from the unique set-up of the 2004 Indian General Elections, I analyse the effect of media forecasts on voting decisions. I model the evolution of voter preferences as a Bayesian updating process in which voters have imperfect information and are uncertain about party quality. I identify the key learning parameters of the structural model by using within-state variation in the sequence of polling across constituencies. In addition, I exploit geographical discontinuities in the polling schedule to compare voting behaviour amongst early and late voters. I find that late voters react to surprises in voting returns and increase, by 20%, their probability of voting for parties that made substantial early-poll gains based on exit poll results. Finally I use the occurrence of simultaneously held state elections as a counterfactual and examine the effect of national exit polls on state voting returns to rule out conformity effects.

Keywords: Learning, Political Economy; **JEL Classification:** C31, O16, G34

1 Introduction

To what extent do exit polls influence voting behaviour? Legislators and election authorities frequently grapple with this question, perhaps now more than ever, as newsrooms adapt to publishing and broadcasting on multiple media platforms during elections[1]. Yet, little evidence is available to shed light on this important issue. A controversial tool, an exit poll is conducted on the day of the election, immediately after voters 'exit' polling booths, enabling media agencies to predict the result of the election. Many observers believe that the premature broadcasting of exit poll results, before all polling has ceased, could influence votes and critically affect election outcomes. In view of this, it is often debated whether the conduct of opinion and exit polls ought to be regulated. In this paper, I contribute to the on-going policy debate and ask if and by how much voters update their voting decisions in response to information conveyed by exit poll forecasts. I exploit a unique feature of the electoral process in India – multi-phase polling of national (general) elections – in a period where exit polls were allowed to be disseminated and broadcast freely, to empirically identify this effect.

The regulation of opinion and exit polls has been an issue of particular importance in India owing precisely to its staggered polling procedure. Elections in India are held, on average, over a period of one month because the logistics of holding elections simultaneously across all areas presents formidable difficulties.

*I thank Toke Aidt, Nageeb Ali, Jane Cooley, Mark D'Souza, Xavier d'Haultfoeuille, Amrita Dhillon, Marcel Fafchamps, Manos Kitsios, Pramila Krishnan, Sanjay Kumar, Eliana La Ferrara, Hamish Low, Imran Rasul, David Rivers and participants at various seminars for helpful comments. I am especially grateful to Sanjay Kumar at CSDS for allowing me to access the NES data and to Sandhya Rao and Sandip Sukhtankar for providing the GIS data. I acknowledge support from the Marie-Curie AMID Research Fellowship.

[†]Center for Research in Economics and Statistics, ENSAE; manasa.patnam@ensae.fr

[1]In 2011, the setting up of a popular website "TweetTheResults.ca" which aimed to aggregate election results in Canada, compelled Canadian election authorities to enact a law, making it illegal to *tweet* election results before all polls have closed, a first of its kind. Reiterating this stance, the French commission for opinion polls recently held that tweets or posts of the results of exit polls in the first-round vote before polling stations have closed is unlawful and could "undermine the integrity of the election" ('*Twitter fera-t-il bugger l'élection*?', Journal du Dimanche, April 7 2012).

Voters go to poll on different dates, depending on the polling phase assigned to their constituency by the Election Commission of India (ECI henceforth). The last few elections have seen polling schedules with four to five distinct phases, approximately set apart by an interval of about one week (see Figure 1). In light of such sequencing, media independence over the release of exit poll forecasts *between* polling phases is often questioned. The ECI shared the view taken by many other political parties that a large majority of the electorate, which has access to information on polls, is likely to be influenced by them whilst making their choice (Abjorensen 2012). Such an outcome, if held true, is considered deeply problematic in many ways. Firstly, it would call into question the order of voting itself. Early voters, by affecting the voting choices of late voters, could disproportionately influence election outcomes. The choice of constituencies selected to poll first will therefore matter crucially. Secondly, exit polls results could induce a 'momentum' effect, whereby a party's performance in early phases has a positive effect on its performance in later phases (Knight and Schiff 2010). This could potentially give undue advantage to parties with early (projected) leads over other parties in later polling rounds.

In recognition of these issues, the ECI, since 1998, has repeatedly sought to enact a blanket ban on the broadcast of exit and opinion polls, until the last phase of voting is over. It was only in 2008 that the Government of India approved the ECI's proposal by amending the Representation of People Act of 1951 and inserting section 126 (A & B) for the "Restriction on publication and dissemination of result of exit polls, etc." (GOI 2009, pp. 2)[2]. As the clause applied singularly to exit poll broadcasting, the Election Commission continues to seek a regulatory ban on opinion polls as well.

While the debate on regulating opinion and exit polls continues to thrive, clear evidence as to whether voters are indeed influenced by poll predictions remains lacking. To make the case that individual reactions are most likely sensitive to such information, coarse inference is usually drawn by observing investor sentiments during election periods, particularly at the time of exit poll announcements. Consider, for example, stock market reactions during the Indian national election held in 2004, where the broadcast of exit polls was unrestricted, to its policy counterfactual, the 2009 elections, where the ban on exit polls was operational. I plot the daily BSE (Bombay Stock Exchange) trading index, the SENSEX, over both these elections highlighting the polling phases (grey blocks[3]) in Figures 3 & 4. In 2004, on the 27^{th} of April, immediately after exit poll results of Phase 2 were announced the SENSEX crashed by 213 points, to reach a three-year low[4]. The index continued to fall, registering another crash at the end of Phase 4, allegedly also in reaction to exit poll forecasts. As a result, it was noted, that the SENSEX fell by almost the same magnitude (twice) *between* these polling periods in response to early election projections, as it did once polling had finished, when results were *officially* announced. Compare now Figure 4 which tracks the SENSEX over the 2009 election period, when exit polls were announced only *after* the last phase of polling. The only point at which the index deviated significantly from its trend, was at the end of the election when results, and exit polls, were finally announced. Barring this, the index remained fairly stable throughout

[2]Section 2.3 and Appendix A.1 discuss in detail the provisions of this Act as well as the legal debate surrounding its constitutional validity.

[3]Each block in grey represents a three day window over the selected polling phase. Exit poll forecasts are usually released and broadcast over this two-three-day window. Stock market crashes are indicated by a dark grey block.

[4]A front-page article in the *Hindu* reported that, "Immediately after exit polls were announced in Phase 2 virtual bloodbath at the stock markets today with the Bombay Stock Exchange Sensitive Index (Sensex) crashing by 213 points at close of trading, the biggest fall since March 13, 2001 when the last stock scam broke out. This time, the fall is attributed to the totally unexpected exit poll results which are pointing towards a hung Parliament.." ('*Exit poll predictions trigger stock market crash*', Hindu, April 28 2004).

the polling period. While the figures provide some compelling evidence, descriptive at best, that investors react to early calls, it is unclear whether individuals change their *voting* behaviour in response to poll projections and by how much.

In this paper, I provide an analytical framework for understanding how exit polls might influence voting behaviour and, using past election data, empirically quantify the magnitude of its effect. The basic social learning model builds on Knight and Schiff (2010) and Larson (2011). In the model, before going to poll, voters have imperfect information[5] and are uncertain about party quality. Under this premise, I show that exit polls, by eliciting and summarizing voting choices of previous round voters, act as a noisy signal that can be used by subsequent voters to inform their own choices. Voters update in a Bayesian manner, by taking a weighted average of their prior beliefs and the recent information delivered by the exit polls. The 'updating' weight, provides a simple measure of the amount by which each voter alters her behaviour; it depends on the noise-to-signal ratio such that voters rely more on precise signals. I estimate this key parameter of the model from aggregate data on actual voting returns using the *within* state variation in the polling sequence of constituencies. Multiple polling rounds within a given state imply that state level preferences, unobservable to the econometrician, are differenced out at each stage, mitigating the bias caused by the presence of correlated effects.

A key insight of the model is that in the learning process voters are not necessarily affected by a party's winning projections, rather they react to surprises in its voting returns. Put differently, parties projected to have secured a lead in the early phases of election do not always gain advantage over other parties in later phases. Bayesian voters will have no reason to revise their choices in favour of the winning party, unless exit poll predictions exceed their own priors. I demonstrate this, and provide support for the aggregate level empirical results by utilizing detailed post-poll data on individual voting behaviour. The objective is to identify how surprises in party performance as projected by exit poll results can affect individual voting choice regardless of its winning prospects[6]. To causally identify this effect, I exploit geographical discontinuities in the polling schedule. I compare voting choices of individuals located on either side of the boundary that was drawn by the ECI to demarcate polling phases across constituencies. Using a semi-parametric regression discontinuity approach, I estimate whether late voters increased their probability of voting for parties that were projected to have made 'gains', relative to early voters who received no such information[7].

In addition to examining the effect of exit polls on voting behaviour, the electoral setting also offers a rich environment with which to study observational learning amongst individuals. Observational learning, a strict subset of social learning, occurs when individual actions are influenced by observing the actions of others. In a recent paper, Cai et al. (2009)[8] point out three aspects of observational learning that

[5]See early work by Alvarez (1998) who discards the assumption of perfect information amongst voters and provides Bayesian foundations for the effect of information on voter uncertainty.

[6]I borrow the term 'surprise' from Moretti (2011) who, in a different context (movie sales), shows how consumers update their priors and learn when they receive news indicating a surprise.

[7]In the absence of information on voter subjective beliefs, I use a concept of 'gains' and 'losses' to proxy for surprises in party performance. In exit poll terminology, a party 'gains' when its predicted performance *improves* relative to its performance from the previous election.

[8]In their paper, the authors conduct a randomized field experiment in Chinese restaurants and show how the information displayed affects the choices of customers. They distinguish learning from a saliency effect by showing that consumers increase their demand for popular dishes only when given a clear ranking of dishes rather than when they were shown a list of popular

distinguish it from other related mechanisms. Firstly, in contrast to social learning that occurs through direct communication, observational learning is not restricted to temporal, spatial or social proximity. A second related aspect is that, observational learning excludes individuals' preferences for conformity. The learning mechanism is absent if individuals choose to adopt choices of other agents because in doing so, they receive a direct utility from conforming. Finally, observational learning is distinct from saliency effects which arise when observing others' choices makes those choices more salient than the alternatives. Thus, observational learning should occur when individuals possess information about each choice in the choice set, rather than possessing information about the choice set as a whole. Below, I take up each one of these issues.

While observational learning can occur independent of social/spatial proximity, it is of interest to understand how these factors can *aid* and accelerate the observational learning process without necessarily involving direct communication[9]. Agents infer more precisely, the informational content of signals received from other agents with whom they share common underlying preferences. The within and across state variation in polling sequence, offered by the setting, allows me to examine this issue. I show in the paper that late voters update more in response to exit poll results received from within-state early voters (state signals), whose state level preferences are known to them, compared to across-state early voters (national signals), whose state level preferences are unobserved to them.

Next, to distinguish the learning from conformity effects *and* to rule out party-state specific shocks or trends *between* polling phases, I make use of an additional source of variation in the 2004 general election structure[10]. In 2004, along with the national elections, five states simultaneously held their legislative assembly (state) elections, polling in accordance with the national election schedule. If voters clearly distinguish between party quality at the national level from the party quality at the state level, then any signal providing information on the former should not affect their voting choices in the state election . To this purpose, I use the exit poll forecast of each party's performance at the national level and test whether voters update in response to such news when they vote at the state level[11]. Failure to do so would provide some indication that the conformity channel is absent. The specification implicitly also allows me to test for the presence of some common state specific trends for each party between polling phases. Finally, the saliency aspect is easily dealt away with because exit polls provide a clear ranking of each party. Voters receive information on the number of seats each party is expected to win, based on voting in earlier phases, rather than a list of parties generally expected to win.

The contribution of this paper is threefold. First, I provide evidence to show that individuals update their voting choices in response to exit poll predictions. To the best of my knowledge, this is the first study to examine the relationship, both theoretically and empirically. Previous research, mainly from the political

dishes. Other papers that find evidence on social/observational learning in different contexts include Munshi (2004), Bandiera and Rasul (2006) (social learning in technology adoption), Bennett et al. (2011) (social learning during epidemics).)

[9] Larson (2011) derives theoretical conditions to differentiate between the rate of learning when private tastes are shrouded relative to when they are completely known. His model is derived in a context where agents learn (observationally) from other peopleŠs past action and is applicable to the setting discussed in the paper. Although idiosyncratic private tastes get 'washed out' in aggregation, they inflate the variance component of the signal.

[10] I allow for state preferences to enter the utility function of the voter directly. Despite this, the worry is that observing actions of previous round voters may induce bandwagon or underdog effects whereby voters derive utility from voting for parties that are (expected to be) leading or lagging behind. Put another way, I do not explicitly allow for the actions of previous round voters (a subset of the entire state) to enter the utility function directly.

[11] Exit poll predictions of state elections are rare and almost never reported.

science literature has focused on the influence of opinion polls prior to the election (McAllister and Studlar 1991)[12] and voter turnout patterns in response to poll predictions (Sudman 1986)[13].

Second, I examine observational learning in the context of sequential elections. In this respect, my paper is most closely related to the work by Knight and Schiff (2010) who examine social learning in the context of US presidential primaries. Using daily polling data from the 2004 presidential primary election, the authors find that later voters strongly updated their priors on candidates based on polling results of early voters. In their model, voters receive public signals in the form of actual voting returns, disclosed after each primary, and update over these. My paper differs from theirs in at least four respects. Firstly, I analyse learning in the context of national (general) elections, which involve the entire electorate, in contrast to the primary system, which, being a competition between members of the same party, involve only a subset of the electorate. Secondly, I relax the common prior assumption and show how social learning can take place when voters have diffuse common priors[14]. I also allow voters to update differentially across parties. In doing so, I am able to examine whether voters update asymmetrically with respect to each party. Thirdly, in my empirical analysis, I account for voter turnout and explore whether the release of exit polls affects the probability of voting itself and how this effect varies across parties. Finally, voters in my model and empirical setting do not observe actual voting returns; instead they receive an exit poll forecast that summarizes the voting patterns in the previous phase. This adds an additional noise element in the form of a poll/media specific finite sampling error. A related paper is by Lee and Moretti (2009) who show that market prices in the political prediction markets are responsive to information released through popular opinion polls. I show in the paper that, indeed, even voters respond to exit poll predictions[15].

Topically, I also contribute to the growing body of research that seeks to understand how the media might influence and help shape political opinion amongst the electorate. Notable papers are Gentzkow, Shapiro, and Sinkinson (2011) & Gentzkow (2006) who show how newspapers and television have a robust positive effect on political participation and Chiang and Knight (2011) who find that media endorsements influence a voter's decisions to support the recommended candidate.

The rest of the paper is organized as follows: Section 2 defines the phase wise election procedure in India. Section 3 discusses the model structure, while Section 4 lays out the empirical framework and the identification strategy. The data used is described in Section 5. Section 6 discusses the results and Section 7 concludes.

[12] McAllister and Studlar (1991) examine the effect of opinion-polls on voting choice in the context of the British General Elections between 1979 and 1987. They find a strong positive correlation between an individual's probability of voting for a party and its projected poll leads. They interpret this as evidence of a 'bandwagon effect' where voters favour a party that is doing well in the polls. The authors however caveat this result by admitting to the possibility that voters who displayed a bandwagon effect may well have voted for the party anyway.

[13] Sudman (1986) reviews literature that attempt to infer a relationship between exit poll predictions and voter turnout using both micro and macro data.

[14] In their model, Knight and Schiff (2010) show learning to occur through the evolution of common priors. The common prior assumption has often been criticized in theoretical work on learning (Cripps et al. 2008; Gul 1998). See Acemoglu and Ozdaglar (2011) who show that *Bayesian* foundations of such common priors are not necessarily very strong.

[15] Recent papers that have theoretically examined issues related to social learning in elections include Ali and Kartik (2011) and Demichelis and Dhillon (2010).

2 Elections and Exit Polls in India

2.1 Electoral System and Alliance Formation

India is administratively divided into 28 states and 7 union territories[16]. It follows a parliamentary form of electoral system with two Houses, the 'Lok Sabha' (Lower House/House of the People) and the 'Rajya Sabha' (Upper House/Council of States), that form the legislature. The House of the People ('Lok Sabha') has 545 members, elected for a five-year term. Each Member of Parliament holds a seat, representing a single electoral constituency from where they are elected under the plurality – 'first past the post' – electoral system. This means that within each constituency the candidate with the highest number of votes - a simple majority - is declared the winner. Electoral constituencies, known as Parliamentary Constituencies (henceforth PC), are explicitly demarcated within each state[17] in a way that is proportional to the size of the electorate[18].

In terms of its political set-up, India's polity is considered to be widely fragmented with many regional and smaller parties finding a fair degree of representation at the national level. Over 200 parties (independents included) have contested fourteen elections so far held in India with no less than 40 parties gaining representation in the parliament (Lahiri and Roy 1984). Although extreme in division, Karandikar et al. (2002) calculate and find that the number of 'effective' parties ranges between five and seven in India and that this figure drops to three when looking at the state level average[19]. Given this lopsided multi-party division, many parties have in the recent past (since the 1998 general elections) tried to form 'alliances' with the view to aggregate votes by sharing the total number of contested seats. Parties that enter into an alliance contest national elections on a joint platform under a common manifesto. Sridharan (2004) analyses the political alliance structure in the 2004 Indian general election with a discussion of each party's seat sharing agreement. I skip the detailed discussion on the composition of each alliance and provide a brief overview of the alliance pattern[20]. In the general elections of 2004, there were two major alliances that fought the elections under pre-election seat-sharing agreements: "Congress" (Centre-Right) Party led United Progressive Alliance (UPA), "BJP" (Right Wing) led National Democratic Alliance (NDA). Left Front parties and parties that chose not to align with either of the alliances are classified under 'None'. The majority of the independent candidates also come under this category. In this paper, I conduct all analysis at the (pre-poll) *alliance* level. I do this, not with the view of making the empirics more tractable, but in order to align the analysis with the information provided by the exit poll results. All exit poll results are aggregated and reported at the alliance level[21]. My analysis, therefore, aggregates voting returns up to the alliance level, in line with the poll forecasts, but the analytical results hold at the disaggregated party level as well.

[16] Union Territories are directly administered by the central government

[17] Each state also elects its own legislative Assembly, whose members are elected directly, also by the plurality system. Each member of the state legislative assembly represents an Assembly Constituency (AC) which is the relevant electoral division for state elections. On average, about 5 to 8 Assembly Constituencies together make one parliamentary constituency.

[18] The state level demarcation of electoral boundaries means that some states are over-represented in parliament at the cost of others, due to uneven state sizes.

[19] The authors calculate an effective number of parties index which provides a count of parties weighted by their relative strength.

[20] I refer the interested reader to Sridharan (2004) and Spary and Wyatt (2006) for detailed analysis on the 2004 Indian general election.

[21] Although in their survey polling agencies elicit voting choices at the party level, the existence and knowledge of pre-poll alliances and seat sharing arrangements naturally leads them to aggregate the results. From this perspective – public knowledge of pre-poll alliances – both sets of analysis, aggregate (alliance) and disaggregate (party) are equivalent.

2.2 Multi-Phase Polling Procedure

Owing to the large spread and size of constituencies across India, elections are conducted in a phase-wise manner. The electorate includes approximately 714 million voters, with votes cast in 828,804 polling stations scattered throughout the country for over 5,000 candidates from seven national political parties and several regional parties. The conduct of election in a phase-wise manner is spread over an average of one month. In 2004, the first phase of elections was held on 20^{th} April 2004 for 141 constituencies while the last phase was conducted on 10^{th} May 2004 for 182 constituencies. Figure 1 shows a map of the polling schedule for this election period. The responsibility of deciding the phases and conducting the elections lies entirely with the ECI which functions as an autonomous, quasi-judicial constitutional body of India. The scheduling of elections in this phase-wise manner is determined by the ECI taking into account several factors.

The ECI, in a press-note issued to announce the dates of the election, reported that many political parties had expressed their concerns over multi-phased polling within each state. While some parties were in favour of holding polls on a single day for each state, other sought multiple polling rounds to enable more effective state administration in order to conduct elections. As a consequence the ECI in deciding polling dates within states, balanced the need to have a one-day poll in one state as far as practicable with the requirements of security forces available (ECI 2004). In addition, the ECI also considered other factors such as “various holidays and festivals during the months of April and May, harvest season in certain parts of the country and the inputs taken from the India Meteorological Department in respect of coming monsoon..”(ECI 2009, pp 2).

2.3 Exit Polls

The duration of the election phases coupled with increasing mass media coverage in India resulted in large number of exit polls being conducted and published/broadcast immediately after each phase of the election was completed regardless of the onset of subsequent phases. Most polling agencies aim to predict the number of seats won by each party in the Parliament (see for example Figure 2). Most agencies follow more or less the same techniques (discussed in Appendix A.1), differing only in their sample draws, variation in estimates across polling agencies tends to be quite low. This means that voters, on average receive similar information regardless of their choice of agency/media-outlet from which they seek to hear exit poll predictions.

The regulation of the conduct of opinion polls and exit polls is a sensitive issue. Since 2008 that the Government of India approved the ECI’s proposal to prohibit the release of exit polls during elections. The amendment however continues to be controversial since the time of its enactment and has been challenged mostly on the grounds that it fundamentally violates constitutional provisions. I discuss the current policy debate on the regulation of exit-poll forecasts and its constitutional implications in Appendix A.1. In what follows, I attempt to throw light on this much debated policy issue. I first lay out the framework for understanding how exit poll results affect voting decisions and then proceed to empirically test the predictions of the model. The next section describes the model.

3 Belief Updating with Exit Polls

I develop a model of Bayesian belief updating for a given election that is exogenously sequenced across a finite set of parliamentary constituencies. The notation is roughly similar to Knight and Schiff (2010). Each constituency is populated by a continuum of voters and the utility that any voter i in constituency c and state s derives from voting for alliance a can be written as:

$$u_{aics} = q_a + \eta_{as} + \psi_{aics} \tag{1}$$

where q_a represents the quality of alliance a, η_{as} is the state-level preference for alliance a and ψ_{aics} is the individual preference for alliance a which is distributed with mean zero and constant variance. The errors, ψ_{aics}, have type-1 extreme value distributions[22] and are independent across individuals. State level preferences are distributed normally and independently across states, with mean zero and variance given by $\sigma^2_{\eta_a}$. I also assume that voters are uncertain about alliance quality[23] and make their decisions to vote in favour of some particular alliance just before going to poll, conditional on receiving a signal (public or private) about quality. Specifically, voters in constituencies that go to poll in the first phase receive a (noisy) *private* signal over the quality of alliance a given by θ_{ac}, but receive no public signal about the same. At each phase t of polling, voters have a prior over alliance quality that is normally distributed with mean $\mu_{a,t}$ and variance $\sigma^2_{a,t}$. Before polling begins, I assume that agent's prior on q_a is diffuse[24]. Private signals, θ_{ac} are given by:

$$\theta_{acs} = q_a + \epsilon_{acs} + \xi_{acs} \tag{2}$$

where ϵ_{acs} ($\epsilon_{acs} \sim N(0, \sigma^2_{\epsilon_a})$) and ξ_{acs} ($\xi_{acs} \sim N(0, \sigma^2_{\xi})$) are the combined noise in the private signal. The idiosyncratic constituency specific random noise is denoted as ϵ_{acs} represents, whereas ξ_{acs} captures the variation in tastes for different alliances amongst constituencies. Parties can identify ξ_{acs} and field candidates in accordance with this taste parameter[25]. I assume that this signal is common across all voters within constituencies but not necessarily within a given state. These signals are unobserved by voters in other constituencies.

The expected utility of voter i in constituency c and state s from an electoral win of alliance a can now be written as:

$$E(u_{aics}|\theta_{acs}, \eta_{as}, \psi_{aics}) = E(q_a|\theta_{acs}) + \eta_{as} + \psi_{aics} \tag{3}$$

[22]The standardized Gumbel or the type-1 extreme value distribution has mode 0, mean 0.577 and variance 1.64. By using this distribution, I implicitly assume that the property of Independence of Irrelevant Alternatives (IIA) holds. Later, I relax this assumption by considering a nested structure when incorporating voter turnout.

[23]q_a is also used to represent voter uncertainty about the ability of each alliance to form a stable coalition government – a major concern amongst voters in India (Rangarajan and Vijai 2002).

[24]I assume diffuse common priors to make the analysis more tractable; however this assumption would not invalidate any of the model's results – see Larson (2011) for a generalization; Green, Gerber, and de Boef (1999) also consider the case where voters have no information about the state of opinion before the first poll. A diffuse common prior would only imply that voters in each constituency are easily influenced by information and other state or constituency specific private signals that they receive prior to polling (Acemoglu and Ozdaglar 2011). This is not an unreasonable assumption to hold in the Indian context given the extreme nature of multi-party fragmentation, the dominance of many state specific regional party and constantly shifting alliance positions.

[25]For example, parties may choose to field male candidates over women, backward *caste* candidates in reserved constituencies or candidates with major criminal convictions in areas where they are popular and have considerable political clout. See Aidt et al. (2011) for an analysis of conditions that determine how parties field candidates with criminal conviction and its effect on electoral outcomes.

I now illustrate social learning in the context of Bayesian updating by considering two rounds of polling. I generalize the framework to incorporate t rounds of polling at the end of this section.

3.1 Phase 1 Constituencies

Voters in constituencies going to poll in Phase 1, i.e. at the very beginning of elections, have no predecessors to observe. They receive only a private signal over quality θ_{ac}. Given their diffuse priors, they form expectations over quality, purely on the private signal, given by:

$$E_1(q_a|\theta_{acs}) = \theta_{acs} \tag{4}$$

Replacing Eq. (4) in Eq. (3), we get:

$$E_1(u_{aics,1}|\theta_{acs}, \eta_{as}, \psi_{aics}) = \theta_{acs} + \eta_{as} + \psi_{aics} \tag{5}$$

Since the errors, ψ_{aics} are type-1 extreme value distributed, the probability that individual i will vote for alliance a is given by:

$$\Pr(E_1(u_{aics,1}) > E_1(u_{kics,1}); \forall k \neq a) = \frac{\exp\left[\theta_{acs} + \eta_{as}\right]}{\sum_{k=0}^{A} \exp\left[\theta_{kcs} + \eta_{ks}\right]} \tag{6}$$

This expression represents the (aggregate) logit model and is equivalent to the vote share of each alliance a in constituency c (Nevo 2000):

$$\upsilon_{acs,1} = \frac{\exp\left[\theta_{acs} + \eta_{as}\right]}{\sum_{k=0}^{A} \exp\left[\theta_{kcs} + \eta_{ks}\right]} \tag{7}$$

I normalize the utility from the baseline alliance (denoted by 0) to zero. The ratio of alliance $a's$ vote shares relative to the baseline alliance is, then, given by:

$$\frac{\upsilon_{acs,1}}{\upsilon_{0cs,1}} = \exp\left[\theta_{acs} + \eta_{as}\right] \tag{8}$$

The log-odds ratio is derived by taking the logarithm of the expression above[26]:

$$\vartheta_{acs,1} = \ln\left(\frac{\upsilon_{acs,1}}{\upsilon_{0cs1}}\right) = \theta_{acs} + \eta_{as} \tag{9}$$

I assume that voters exhibit sincere voting behaviour. This is a commonly held assumption which ensures that, in equilibrium, the behaviour of individual voters has no effect on overall vote shares and hence on the behaviours of later voters. The sincere voting assumption and the specified normalization imply that constituency level vote shares are equal to the aggregate proportional utility of voters for any given alliance a:

$$\vartheta_{acs,1} = \theta_{acs} + \eta_{as} \tag{10}$$

[26]Note that while the IIA property is assumed for individual behaviour, it is not an aggregate property. Alvarez and Nagler (1998) show that it is possible for IIA to hold for individuals but to be violated in the aggregate.

3.2 Exit Polls: State-Wise Public Signals

Immediately after polling is completed in Phase 1, the media release a forecast of *expected seats* for each alliance in each state. Each media outlet m estimates constituency level vote shares on the basis of finite sample, with some forecasting error κ_{sm}. Similar to Lee and Moretti (2009), the distributional property of this error is akin to that of small-sample noise. The finite sample error, $\kappa_{sm} \sim N(0, \sigma^2_\kappa)$ is unbiased[27] and has a variance σ^2_κ that reflects consensus amongst various exit polls forecasts. The precision of the public signal is, thus, smaller when there is a large amount of variance in exit poll predictions across M media sources. Media forecast (denoted by superscript *) of vote shares is given by:

$$\vartheta^*_{acs,1} = \vartheta_{acs,1} + \kappa_{sm} \tag{11}$$

$$\vartheta^*_{acs,1} = \theta_{acs} + \eta_{as} + \kappa_{sm} \tag{12}$$

The expected number of seats is calculated and broadcast by the media, based on the above forecast of vote shares. In Appendix A.2, I show how the expected number of seats can be mapped onto a given distribution of forecasted vote shares for each state in a multiparty/multi-alliance set up. Briefly, the process involves discounting the projected seat shares, π^*_{as}, using state-level swing factors, ρ_s, such that the exit poll signals are able to provide information on the underlying vote shares. For instance, the seat share forecast based on voters in Phase 1 constituencies is given by:

$$\begin{aligned}
\omega^*_{asm,1} = \frac{1}{\rho_s} \ln \frac{\pi^*_{as,1}}{\pi^*_{0s,1}} &= \vartheta^*_{as,1} && (13)\\
\omega^*_{asm,1} &= q_a + \xi_{as} + \eta_{as} + \kappa_{sm} && (14)\\
\omega^*_{asm,1} - \eta_{as} &= q_a + \xi_{as} + \kappa_{sm}
\end{aligned}$$

The transformed seat forecast gives an unbiased estimate of q_a and is normally distributed with variance given by, $\sigma^2_{\omega^*,1} = \frac{\sigma^2_\xi}{C_1} + \sigma^2_\kappa$. Because voters receive exit poll forecasts from M media-sources, the public signal is averaged over all these independent signals. The averaged out transformed public signal is given by $\bar{\omega}^*_{as1} - \eta_{as}$ where $\bar{\omega}^*_{as,1} = \frac{1}{M}\sum_{m=1}^{M}\left(\omega^*_{asm,1}\right)$[28].

3.3 Phase 2 Constituencies

Voters in constituencies going to poll in the next period, that is in Phase 2, receive a state level public signal ω_{as}, over and above their private signals. In the absence of these public signals, their vote shares would have mirrored the vote share distribution of Phase 1 constituencies as given by Eq. (10). I can rewrite this for Phase 2 constituencies as:

[27] I assume away media specific bias in *exit poll forecasts*. Most media agencies that conduct exit polls disclose, in detail, their methodology as well as sample sizes. Releasing incorrect and biased exit poll forecasts are generally not to the advantage of any media outlet as they face a serious threat to their reputation in the event that they are proven wrong. Most news agencies in fact compete on getting 'the results right'. The media however may continue to use other channels, such as endorsements, to introduce alliance specific bias amongst their viewers. It is beyond the scope of the paper to model the joint interaction of such news (if any) on belief updating.

[28] The variance of this average signal is, $\sigma^2_{\bar{\omega}^*,1} = \frac{\sigma^2_\xi}{C_1} + \frac{\sigma^2_\kappa}{M}$.

$$\vartheta_{acs,2} = \theta_{acs} + \eta_{as} \tag{15}$$

Under social learning, voters would update their previous period priors conditional on receiving this public signal as follows:

$$\begin{aligned} E_2(u_{aics,2}|\theta_{acs}, \eta_{as}, \psi_{aics}, \bar{\omega}^*_{as1}) &= \beta_{a,2}(\theta_{acs}) + (1-\beta_{a,2})(\bar{\omega}^*_{as1} - \eta_{as}) + \psi_{aics} + \eta_{as} \\ \Pr(E_2(u_{aics,2}) > E_2(u_{ikcs,2}); \forall k \neq a) &= \frac{\exp\left[(1-\beta_{a,2})((\underbrace{\bar{\omega}^*_{as,1} - \theta_{acs}}) - \eta_{as}) + \theta_{acs} + \eta_{as}\right]}{\sum_{k=0}^{A} \exp\left[\beta_{k,2}(\theta_{kcs}) + (1-\beta_{k,2})(\bar{\omega}^*_{ks,1} - \eta_{ks}) + \eta_{ks}\right]} \end{aligned} \tag{16}$$

The term in the under-bracket, $\bar{\omega}^*_{as1} - \theta_{acs}$, measures the deviation of the prior from the information received or the extent of 'surprise'. If voters are Bayesian and update their beliefs ($\beta_a < 1$), then under social learning, a positive surprise in favour of alliance a, increases its probability of being selected by voters[29]. At the constituency level, this means that alliances whose forecast exceeds prior expectations are likely to experience an increase in their vote shares. To see this, I use Eq. (16) to derive vote shares of second round constituencies as:

$$\vartheta_{acs,2|\bar{\omega}^*_{as,1}} = \beta_{a,2}(\theta_{acs}) + (1-\beta_{a,2})(\bar{\omega}^*_{as,1} - \eta_{as}) + \eta_{as} \tag{17}$$

Further, we can replace the priors for constituencies in Phase 2 (given by Eq. (15)) in Eq. (17):

$$\vartheta_{acs,2|\bar{\omega}^*_{as1}} = \beta_{a,2}\vartheta_{acs,2} + (1-\beta_{a,2})\bar{\omega}^*_{as,1} + (1-\beta_{a,2})\eta_{as} - (1-\beta_{a,2})\eta_{as} \tag{18}$$

Note that the common state level preferences are absorbed, giving us:

$$\vartheta_{acs,2|\bar{\omega}^*_{as1}} = \beta_{a,2}\vartheta_{acs,2} + (1-\beta_{a,2})\bar{\omega}^*_{as,1} \tag{19}$$

where $\beta_{a,2}$ is given by:

$$\beta_{a,2} = \frac{M\sigma^2_{\xi} + C_1\sigma^2_{\kappa}}{M\sigma^2_{\xi} + C_1\sigma^2_{\kappa} + MC_1\sigma^2_{a,2}} \tag{20}$$

where $\sigma^2_{a,2} = \sigma^2_{\xi} + \sigma^2_{\epsilon_a}$ is the combined variance of the noise in the prior, which for second round constituencies is simply their private signal. It is easy to see that the weight on the public signal, $(1-\beta_a)$, increases when the precision of the public signal (average exit poll forecast), $\frac{1}{\sigma^2_{\kappa}}$ increases[30]. This weight also increases as the number of constituencies polling in the previous round C and the number of media sources M releasing exit poll forecasts, increase. Finally note that the belief updating weights are alliance specific. This means that voters may choose to update asymmetrically across alliances. Voters place less weight on the information contained in exit polls for any particular alliance when they hold strong priors over its quality.

[29] I show below that the common state level preferences will be absorbed by state level exit poll forecasts.

[30] This can be seen by breaking down the expression: $\beta_{a,2} = \frac{\frac{1}{\sigma^2_{\theta_a}}}{\frac{1}{\sigma^2_{\theta_a}} + \frac{1}{\frac{\sigma^2_{\xi}}{C} + \frac{\sigma^2_{\kappa}}{M}}}$

The model can be iterated over t rounds of polling to give:

$$\vartheta_{acs,t|\bar{\omega}^*_{as,t-1}} = \beta_{a,t}\vartheta_{acs,t} + (1-\beta_{a,t})\bar{\omega}^*_{as,t-1} \tag{21}$$

with $\beta_{a,t}$ in each period given by:

$$\beta_{a,t} = \frac{\sigma^2_{\bar{\omega}^*,t-1}}{\sigma^2_{a,t} + \sigma^2_{\bar{\omega}^*,t-1}} \tag{22}$$

where $\sigma^2_{a,t}$ is the variance of the prior in period t and $\sigma^2_{\bar{\omega}^*,t-1}$ is the variance of the public signal. The variance of the public signal will evolve according to:

$$\sigma^2_{\bar{\omega}^*,t} = \frac{1}{C_T^2}\left[\left((1-\beta_{a,t})^2\sigma^2_{\bar{\omega}^*,t-1} + (\beta_{a,t})^2\left(\frac{\sigma^2_\xi}{C_t}\right) + \frac{\sigma^2_\kappa}{M}\right)C_t^2 + \sigma^2_{\bar{\omega}^*,t-1}(C_{T-1})^2\right] \tag{23}$$

where $C_T = \sum_{s=1}^{t} C_{s-1}$ denotes the total number of consistencies that have finished polling up-till time t.

3.4 Exit Polls: National Signals

So far, I have considered the case where voters update over exit poll forecasts from within their own states. I now add national signals to the updating model to consider how voters might choose to update over signals from other states. The national signal is simply an aggregation of state level signals from all other states that have gone to poll at time t. These signals contain an additional noise element in the form of state level preferences, η_{as}, that voters in states other than their own cannot observe. The combined signal over $S_t - 1$ states that have polled at time t is therefore:

$$\begin{aligned}
\frac{1}{S_t-1}\sum_{s=1}^{S_t-1}\bar{\omega}^*_{as,t} &= \frac{1}{S_t-1}\sum_{s=1}^{S_t-1}\left(q_a + \xi_{as} + \eta_{as} + \frac{1}{M}\sum_{m=1}^{M}\kappa_{sm}\right) \\
\bar{\omega}^{N*}_{a,t} &= q_a + \frac{1}{S_t-1}\sum_{s=1}^{S_t-1}(\xi_{as} + \eta_{as}) + \frac{1}{M}\sum_{m=1}^{M}\kappa_{sm}
\end{aligned} \tag{24}$$

The updating equation can now be modified to accommodate additional weights for national signals as:

$$\vartheta_{acs,t|\bar{\omega}^*_{as,t-1},\bar{\omega}^{N*}_{a,t-1}} = \beta_{a,t}\vartheta_{acs,t} + (1-\beta_{a,t}-\gamma_{a,t})\bar{\omega}^*_{as,t-1} + \gamma_{a,t}(\bar{\omega}^{N*}_{a,t-1} + \eta_{as}) \tag{25}$$

Note that at any given time, $\sigma^2_{\bar{\omega}^*,t} < \sigma^2_{\bar{\omega}^{*N},t}$, since national signals contain additional noise from (aggregate) unobserved state preferences.

4 Empirical Specification and Identification

There are two important features of the model that lend themselves to empirical analysis in the context of sequential elections. Firstly, belief updating weights are heterogeneous; they depend on inherited priors from the previous round of polling. Using within-state constituency level vote share dynamics across four rounds of polling, I estimate alliance-specific belief updating weights to test whether there is asymmetric updating. Secondly, Eq. (16) implies that if there is social learning and some updating of beliefs, then individual voters are more likely to vote for a party following a positive surprise due to an exit poll forecast from the previous round of polling. Using detailed individual level data and exploiting within state geographical discontinuities in phase-wise polling boundaries, I estimate the marginal effect of a positive surprise on individual voting behaviour.

I use both aggregate and individual data to conduct the empirical analysis for the following reason. The analysis of aggregate voting returns data is useful mainly because it absorbs idiosyncratic individual level preferences, unobservable to the econometrician, and allows for the social learning parameters to be estimated directly. However, the structural assumptions implied by the model, such as exogenous sequence selection and common state level preferences, remain quite restrictive. To allow for greater flexibility in the modelling assumptions and to ensure stronger identification, I *complement* the empirical analysis by using individual level data on voting choice. By adopting a regression discontinuity approach, I tackle, to some extent, the issue of non-random assignment of constituencies into their respective polling sequence. Additionally, common preferences amongst voters can be disaggregated, semi-parametrically, to a level further than the state, since the design essentially compares voters residing in close proximity to each other.

4.1 Constituency Level Belief Updating Weights

In this section, I outline the procedure used for estimating updating weights. The procedure involves computing the swing factors, ρ_s for each constituency, obtaining priors for later polling constituencies and, finally, estimating the belief updating weights as given by Eq. (19).

4.1.1 State Swing Factors

I first estimate the swing factors for each state based on results from the previous elections. There have been various techniques proposed for estimating swing factors, most of them applicable to two-party settings that require a time series of election data for identification. Using a time series of election data is not feasible for this analysis for two reasons. Firstly, contesting elections under pre-poll alliance arrangements is a relatively recent phenomenon. Secondly, alliance formation is not stable; as pointed out by Linzer (2012), same parties do not necessarily contest for the same seats in repeated years and this problem is exacerbated under alliance-specific seat sharing arrangements.

To tackle these challenges, I use a method proposed by Linzer (2012) to estimate swing factors under a multi-party setting using data from a single election. I use data from the 1999 general election but group parties according to their 2004 alliances. I set aside some smaller parties and independents and their exclusion does not affect the analysis. I then apply the method as suggested by Linzer (2012). Briefly, the method involves estimating a density function to capture the underlying joint distribution of alliance vote shares and number of votes cast across constituencies within a state. Using the proportion of state-level votes actually won, as well as the seat-share distribution, one can calculate average per-unit change in seat share around the total party vote. This quantity provides an estimate of the swing factors[31]. Monte Carlo

[31] As in Gelman and King (1994), the swing factor can be estimated using Bayesian procedures as: $\rho_s = \frac{E(\tilde{S}_{as}|\bar{V}_{as}+0.01)-E(\tilde{S}_{as}|\bar{V}_{as}-0.01)}{0.02}$, where $\bar{V}_{as}$ is the average state level vote for alliance a and $\tilde{S}_{as}$ is the potential seat share. This pro-

simulation methods are then employed to generate hypothetical election outcomes that are consistent with the density estimate. This process is replicated a number of times to derive each alliance's swing factor. Although the procedure estimates state-wise swing ratios for each specific alliance, I average these over all alliances to arrive at a common state-specific swing factor in order to be consistent with the model above. Finally, I use these swing ratios to normalize state level seat share forecasts to arrive at an estimate of $\bar{\omega}^*_{as,t-1}$.

4.1.2 Belief Updating Weights

Using the normalized state level seat share forecasts I estimate the belief updating weights for state signals and national signals using the equations derived in Sections 3.3 & 3.4:

$$\vartheta_{acs,t|\bar{\omega}^*_{as,t-1}} = \beta_{a,t}\vartheta_{acs,t} + (1-\beta_{a,t})\bar{\omega}^*_{as,t-1}$$

$$\vartheta_{acs,t|\bar{\omega}^*_{as,t-1},\bar{\omega}^{N*}_{a,t-1}} = \beta_{a,t}\vartheta_{acs,t} + (1-\beta_{a,t}-\gamma_{a,t})\bar{\omega}^*_{as,t-1} + \gamma_{a,t}(\bar{\omega}^{N*}_{a,t-1} + \eta_{as})$$

Estimation of the above equation would be straightforward if we had vote share data on every alliance in every constituency both before receiving the forecast and after. However, this is not possible when using data on actual election outcomes as no constituency (excluding bye-elections) goes to poll twice. Therefore I am unable to observe an actual counterfactual.

The timing of sequential elections allows me to construct a quasi-counterfactual. This is done by estimating the priors or the realized vote shares for those constituencies that did not ever receive an exit poll based forecast. For these constituencies their observed vote shares correspond to the unconditional priors. I can, therefore, estimate the parameters on both constituency and alliance specific characteristics that can predict the observed vote shares for the control group constituencies. The key point is that the priors of later polling constituencies are predicted with reference to constituencies going to poll in Phase 1 within their own state. This means that common state level preference for any alliance (as modelled in Equation 10) will be accounted for and will feature as part of the conditioning prior. Identification would not be possible if all states had constituencies that went to poll in the same phase because state level preferences would be unobservable to the econometrician and would be correlated with both the prior and the forecast.

Adapting a framework from Lybbert et al. (2007)[32], I estimate the belief updating weights in two states. In the first step I specify for each alliance the following regression equation and estimate it over the subset of phase 1 constituencies, i.e., those that did not receive a forecast:

$$\vartheta_{acs} = d_0 + d_1 C_{cs} + d_2 A_{acs} + \gamma_s + \epsilon_{acs} \tag{26}$$

cedure estimates the most typical case amongst the range of plausible seat swings for each alliance given a one percent change in their vote shares.

[32]Lybbert et al. (2007) propose a test to see whether pastoralists in southern Ethiopia and northern Kenya update their expectations in response to forecast information. They utilize information on subjective beliefs on rainfall from survey data and find a substantial effect of receiving a forecast on these probabilities. In their model, each farmer holds a subjective probability belief for each of the three possible states of nature. For those farmers who receive a forecast, these probabilities can be expressed as posterior beliefs conditional on the forecast. In the absence of information on prior subjective beliefs of farmers, they use predicted values of prior beliefs based on a linear regression with region and individual specific determinants.

where C_{cs} is a vector of constituency specific factors (demographics, average turnout etc.), A_{acs} is a vector of alliance specific factors that varies over constituencies, such as historical electoral performance of a particular alliance in each constituency. Finally, γ_s are state fixed effects capturing common state level preferences for each alliance. Using the parameters from the estimated regression, I predict the vote share received by each alliance for the full sample, denoted by $\hat{\vartheta}_{acs}$. I can then estimate the belief updating weights using the following regression specification:

$$\underbrace{\vartheta_{acs,t|\bar{\omega}^*_{as,t-1}}}_{\textbf{PO}} = \beta_{a,t} \underbrace{\widehat{\vartheta}_{acs,t}}_{\textbf{PP}} + (1-\beta_{a,t}) \underbrace{\bar{\omega}^*_{as,t-1}}_{\textbf{F}} \tag{27}$$

where **PO** is the Posterior, **F** is the Forecast and **PP** is the Predicted Prior.

4.1.3 Within Alliance Asymmetric Updating

The above sub-section demonstrates how voters would update their preferences for a particular alliance upon receiving a forecast. The updating process could possibly be asymmetric between alliances. This means that voters choose to update their preferences for some alliances based on thier projected performance in the previous rounds and fail to so for other alliances. Further, it is possible that voters choose to update asymmetrically even within their preferences for a specific alliance. This would be the case if, for instance, voters revise their beliefs upwards when the state level forecast for a particular alliance exceeds their prior expectation, but fail to do so conversely, i.e., if forecasts are below their priors. In other words, voters would learn and update only if the forecasts exceed their expectations but dismiss information from forecasts that project alliance performance to be below their prior expectations. Learning in this context does not require the alliance to have won in the previous rounds (or get a projected lead) – it only requires that each alliance performs slightly better than what voters expect from them. To test for asymmetric updating from within alliance performance, I modify the updating equation (Eq. (27)) as follows:

$$\underbrace{\vartheta_{acs,t|\bar{\omega}^*_{as,t-1}}}_{\textbf{PO}} - \underbrace{\bar{\omega}^*_{as,t-1}}_{\textbf{F}} = \beta_a (\underbrace{\widehat{\vartheta}_{acs,t}}_{\textbf{PP}} - \underbrace{\bar{\omega}^*_{as,t-1}}_{\textbf{F}}) + \nu_{acs} \tag{28}$$

$$\Delta_{acs,t|\bar{\omega}^*_{as,t-1}} = \beta_a(\Delta_{acs,t}) \tag{29}$$

$$|\Delta_{acs,t|\bar{\omega}^*_{as,t-1}}| = \beta_a^{\uparrow}(|\Delta^{\uparrow}_{acs,t}|) + \beta^{\downarrow}_{acs,t}(|\Delta^{\downarrow}_{acs,t}|) + \nu_{acs} \tag{30}$$

where $\Delta^{\uparrow}_{acs,t}$ measures 'good news':

$$\Delta^{\uparrow}_{acs,t} = \begin{cases} \Delta_{acs,t} & \text{if } \widehat{\vartheta}_{acs,t} < \bar{\omega}^*_{as} \\ 0 & \text{otherwise} \end{cases} \tag{31}$$

and $\Delta^{\downarrow}_{acs,t}$ measures 'bad news':

$$\Delta^{\downarrow}_{acs,t} = \begin{cases} 0 & \text{if } \widehat{\vartheta}_{acs,t} < \bar{\omega}^*_{as} \\ \Delta_{acs,t} & \text{otherwise} \end{cases} \tag{32}$$

4.1.4 Estimation

There are two features of multiparty election data that must be taken into account whilst estimating regression parameters. Firstly all estimated parameters must reflect the constraint that the sum of vote shares across all parties (alliance in this case) must add up to one. Secondly, each alliance's error terms will tend to be correlated with each other because higher vote share for one party must necessarily mean a lower vote share for others. Therefore estimating $A-1$ separate equations via OLS would result in a loss of efficiency. To tackle these two issues Tomz, Tucker, and Wittenberg (2002) develop a methodology that involves transforming the vote shares into logit scales and estimating the resulting parameters using a seemingly unrelated regression (SUR) (Zellner 1962). The model developed above naturally lends itself to estimation using this method since the dependent variable in the empirical specification, i.e., log-odds of each alliance's vote share, is scaled to reflect the recommended transformation[33]. Using this, I specify ν_{acs} to be multivariate normal with covariance given by $\Phi \equiv E[\nu\nu'|\Delta]$. Since the parameters of Eq. (21) need to sum up to one, I add constraints and estimate a system of nonlinear equations by feasible generalized nonlinear least squares.

Given the two-step nature of the estimation procedure, I correct for standard errors using a non-parametric block bootstrap method. Where necessary, I account for the clustering of errors. In general, I report bootstrapped confidence intervals but also compute percentile based confidence intervals for procedures that involve taking transformations of the parameters.

4.2 Sequential Discontinuities in Polling & Voting Behaviour of Individuals

I now describe how individual level voting behaviour is influenced by exit poll forecasts. I show how within-state phase-wise geographical boundary discontinuities can be exploited to identify and estimate the marginal effect of a positive surprise on the voter's choice probabilities.

4.2.1 Identification

To identify the effect of exit polls on individual voting choices, I use post-poll responses of individual voters and compare the voting choices of voters who receive exit poll forecasts to voters that do not. Ideally, if the survey elicited subjective choice probabilities, then I could replicate the framework of the constituency level analysis at the individual level. Unfortunately this data is lacking since the post-poll survey was carried out at the polling booth immediately after voters had finished casting their ballot. This being, it recorded the actual voting choice made by the individual. In the absence of subjective beliefs, I use the following measure of surprise to proxy for the deviation of the prior from the forecast.

All exit poll results reported the expected 'gains' and 'losses' to any alliance, along-with its seat projections. This was done by comparing the seats won by that alliance in the previous national elections (1999) to the projected wins in the current election. A projected 'gain' for any alliance, therefore, represent

[33] A common problem in compositional data analysis is the presence of zero vote shares caused, mainly, due to parties choosing not to contest the election or non-existence of parties in trending data. In my data this is not a severe problem, however there are a few cases where the base alliance "Others" has registered zero votes. Multiple solutions have been offered to deal with this problem. I follow a zero replacement technique and add a small value, one to all percentage vote shares (or 0.01 to fractions), whilst preserving the log-odds ratio. The IIA property implies that the odds ratio is invariant to additions or deletions to the choice set.

an *improvement* in performance relative to that of the previous election. It is important to note that these gains and losses are calculated only in terms of actual seats won or lost rather than increases or decreases in overall vote share. Further the binary media-provided measure of gains/losses tallies well even with current, pre-election, voter expectations as measured by opinion polls[34]. Using this information and summarizing the results from all exit poll reports for every alliance, I calculate state-wise average expected gains and losses. I classify a state as a 'gain state' for alliance a if exit polls predict major gains for alliance a in that state. Alliance a receives a positive surprise in terms of its performance in such states. Subsequently, I compare the voting behaviour of voters following the announcement of this positive surprise to the behaviour of voters just prior to the announcement. I expect that conditional on a positive alliance-specific belief updating weight, voters should increase their expectations and vote in favour of an alliance receiving a positive surprise. To test for this effect consider comparing Eq. (16), with slight abuse of notation, between constituencies that poll first and those that poll second:

$$\begin{aligned}
\Pr(E_1(u_{aics}) > E_1(u_{ikcs}); \forall k \neq a) &= F(\theta_{acs} + \eta_{as} + \psi_{aics}) &(33)\\
\Pr(E_2(u_{aics}) > E_2(u_{ikcs}); \forall k \neq a) &= F((1-\beta_a)((\bar{\omega}^*_{as,t-1} - \theta_{acs}) - \eta_{as}) + \theta_{acs} + \eta_{as} + \psi_{aics}) \\
&= F((1-\beta_a)(\bar{\omega}^*_{as,1} - \theta_{acs}) + \theta_{acs} + \beta_a\eta_{as} + \psi_{aics}) \\
&= F((1-\beta_a)(T_{acs}) + \theta_{acs} + \beta_a\eta_{as} + \psi_{aics}) &(34)
\end{aligned}$$

where I use T_{acs} as a binary indicator of a positive surprise for alliance a. As mentioned above I use a measure of seat gain as an indicator for a positive surprise. In states where alliance a received a positive surprise, the only difference between constituencies is their order of polling. The positive surprise variable (henceforth treatment) takes the value zero for constituencies polling first since they received no forecast and one for constituencies that received the news of a positive surprise.

Stacking the two equations, I obtain the following cross sectional specification:

$$V_{aics} = \tau_a T_{acs} + \gamma_s + \zeta_{aics} \tag{35}$$

where V_{aics} is a binary indicator[35] for whether individual i voted for alliance a in polling round t. T_{as} is the treatment indicator as defined before. In the cross-section (of constituencies) $T_{as} = 0$ for constituencies that voted first and received no forecast. The parameter on common (and unobserved) state level preferences, β_a, can be absorbed by the state level fixed effect γ_s. The inherited prior, θ_{acs}, remains unobserved to the econometrician and becomes part of the composite error term ζ_{aics}. OLS estimates of Eq. (33) would return biased parameter estimates because the error term is correlated with the positive surprise (henceforth 'treatment') indicator due to the omitted prior.

In order to derive causal estimates, I make use of geographic discontinuity in the phasing of polls within

[34]The dummy variable capturing gains and losses is the same when using popular opinion polls as a benchmark instead of previous elections. The absolute deviation of alliance performance from prior expectations however differs between the two measures.

[35]A binary logit is preferred to a multinomial logit specification for ease of interpretation, since the latter requires the use of a base category. I, nevertheless use the multinomial logit specification to simulate counterfactual vote share distributions, through which it is easier to interpret the magnitude and direction of the estimated parameters. Results are robust to the use of either specification.

each state and employ a regression discontinuity method to identify the parameters of interest. For individual voters, assignment to polling phase was determined entirely on the basis of their location. I define a variable D_{ics} as the distance to the geographic boundary d that splits constituencies into different polling phases.

$$T_{acs} = \mathbb{1}_{[D_{ics} \geq d]} \tag{36}$$

The average causal effect of the treatment at the discontinuity point is then given by (Imbens and Lemieux 2008):

$$\tau_a = \lim_{g \to d^+} \mathbb{E}[V_{aics}|D_{ics} = g] - \lim_{g \to d^-} \mathbb{E}[V_{aics}|D_{ics} = g] = \mathbb{E}[V_{aics}(1) - V_{aics}(0)|D_{ics} = d] \tag{37}$$

An important feature to note in the above-mentioned design is that the discontinuity is geographical, i.e., it separates individuals in different location based on a threshold along a given *distance boundary*. Using Eq. (37) to estimate the causal effect would ignore the two-dimensional spatial aspect of the discontinuity. This is because the *boundary line* can be viewed as a collection of many points over the entire distance spanned by the boundary. An individual located north-west of the boundary is not directly comparable to an individual located south-east of the boundary. For the comparison to be accurate, each 'treatment' individual must be matched with 'control' individuals who are in close proximity to their own location *and* the boundary line.

I address this issue in two ways. Firstly, I divide the boundary for each state into a collection of points defined by latitude and longitude spaced at equal intervals of 7 kilo-meters. I also classify the entire boundary line into distinct segments based on line specific 'turn angles' identified by the geo-coding tool. Then utilizing the spatial point co-ordinates of each location I use polynomials and interactions of both latitude and longitude coordinates. I also measure the distance of each AC to the boundary and include polynomials of distance and its interactions with the treatment variable . I condition on line-segment fixed effects, in all the specifications using polynomials of latitude-longitude/distance, such that only AC's within close proximity of each other are compared[36].

Secondly, I make use of location to boundary point distances for every point along the boundary line and estimate separate treatment effects corresponding to each one of them. Such a design inherently incorporates heterogeneity in treatment effects since it identifies an infinite-dimensional curve of treatment effects with specific geographic locations. In a recent paper, Keele and Titiunik (2011) derive identification conditions based on a geographical discontinuity where the score, D is defined as a function of two points, its latitude and longitude. Using a scalar function of the two-dimensional score and with a slight modification of the continuity assumption, the authors show that the treatment effect identified by the GD design at a boundary point (d1, d2) can be identified as:

[36] See Black (1999) who first discussed the use of the boundary segments in a regression discontinuity framework. For a recent application, see Dell (2010) who extends the approach to incorporate a semi-parametric regression discontinuity design.

$$
\begin{aligned}
\tau_a(d1,d2) &= \lim_{(g1,g2)\in G_T \to (d1,d2)} \mathbb{E}[V_{aics}|(D_{ics1},D_{ics2})=(g1,g2)] \\
&\quad - \lim_{(g1,g2)\in G_C \to (d1,d2)} \mathbb{E}[V_{aics}|(D_{ics1},D_{ics2})=(g1,g2)] \\
&= \mathbb{E}[V_{aics}(1)-V_{aics}(0)|(D_{ics1},D_{ics2})=(d1,d2)] \qquad (38)
\end{aligned}
$$

This equation computes the average treatment effect, $\tau_a(d1,d2)$ for every point $(d1,d2)$ on the boundary line. Identifying assumptions are exactly similar to those of the conventional regression discontinuity design.

4.2.2 Estimation

The average treatment effect can be estimated using local linear regression, by including polynomials of distance to the boundary (controlling for line segment fixed effects) in Eq. (35).

$$
\min_{\lambda_a,\tau_a,\phi_a} \sum_{1=1}^{N} \left[V_{aics} - \lambda_a f(D_{ics}-d) - \tau_a T_{acs} - \phi_a T_{acs}\cdot f(D_{ics}-d)\right]^2 \cdot w_{ics}^{RDD} \qquad (39)
$$

The equation is estimated by choosing a bandwidth $w_{ics}^{RDD} = \frac{1}{h}K\frac{f(D_{ics})}{h}$ where K is chosen to be a triangular kernel function, $K(z) = \mathbb{1}_{[|z|\leq 1]}(1-|z|)$ such that it would fit linear regression functions to the observations within a distance h on either side of the discontinuity point. To estimate discontiuities at every point in the boundary, as implied by Eq. (38), I use the estimation technique suggested by Keele and Titiunik (2011) and apply Eq. (39) to estimate the average treatment effect at each such point along the boundary. The technique is similar to that of a geographically weighted regression (Fotheringham, Brunsdon, and Charlton 2002), where the weights are now conditioned on each point location $(d1,d2)$ relative to the locations of all sample individuals within a given state:

$$
\begin{aligned}
\min_{\lambda_{a_{d1d2}},\tau_{a_{d1d2}},\phi_{a_{d1d2}}} \quad & \sum_{1=1}^{N}[V_{aics} - \lambda_{a_{d1d2}}(f(D_{ics1},D_{ics2}) - f(d1,d2)) - \tau_{a_{d1d2}}T_{acs} \\
& -\phi_{a_{d1d2}}T_{acs}\cdot(f(D_{ics1},D_{ics2}) - f(d1,d2))]^2 \cdot w_{ics}^{GDD} \qquad (40)
\end{aligned}
$$

where the weights are now boundary point specific, $w_{ics}^{GDD} = \frac{1}{h}K\frac{f(D_{ics1},D_{ics2})}{h}$.

One problem that arises with estimating a local geographically weighted regression of this form is how to determine the overall significance level for parameters. Since there are on average, over 150 different spatial locations within each state, there will be up to 150 hypotheses to be tested which defines a very high order multiple inference problem. Byrne et al. (2009) show that when the tests being carried out are highly correlated, using a Bonferroni correction results in highly conservative estimates. They propose, instead, a False Discovery Rate (FDR) procedure (Benjamini and Hochberg 1995) that calculates the expected percent of false predictions in the set of predictions. This procedure adjusts the standard p-values from each hypothesis test and presents, for each corresponding p-value, the minimum uncorrected p-value threshold for which that p-value would be in the false discovery set. The adjusted p-vales are usually referred to as ‘q’ values. I use a modified method of the FDR developed by Benjamini and Yekutieli (2001) that also controls for positive regression dependency on each of the test statistics corresponding to the true null hypotheses.

5 Data

To estimate the belief updating weights at the constituency level, I use data on electoral outcomes for each constituency as provided by the Election Commission of India and combine it with data on exit polls reports from newspaper and magazine archives. The Election Commission provides data on party-wise election results for all general elections and state assembly elections held in India from the period 1976-2009. This data contains information on characteristics of each party's candidate contesting the election (gender, caste category) as well as the total votes received by each candidate. Using information on pre-poll alliances and seat-sharing arrangements from the 2004 general election, I aggregate party-level vote shares into alliance specific vote shares. I also use archived data from the Election Commission website to construct variables on past performance of each alliance at the national and state level. At the national level, I simply aggregate vote share data for each party in each election year by their alliance affiliation in 2004. To summarize past performance at the state level I follow the same process but aggregate one step further. Since PC's are a collection of many AC's, I map each AC onto their counterpart PC and then aggregate alliance vote shares up to the PC level. Note that I also use election results from 1999 general election to estimate swing ratios for each state.

For data on seat-share forecasts, I use exit poll results of five news agencies – NDTV-Indian Express, Sahara, Aaj Tak, Zee News and Star TV – as reported in Outlook magazine and Times of India. Each exit poll provided information on each alliance's expected seat share at the state and national level as well as their gains and losses from the previous election.

Finally, I use individual level data on voting outcomes from the National Election Study in 2004 conducted by the CSDS. The survey interviews respondents immediately after polling and enumerates information on the political behaviour, opinion and attitudes of voters alongside their demographics. The survey uses a dummy ballot box for capturing the respondent's voting choice wherein respondents were asked to mark their voting preference on a dummy ballot paper and drop it in a dummy ballot box. Sampling for the survey is carried out using a multi-stage stratified random sampling design. The first stage involves stratified sampling of Assembly Constituencies by state proportional to their size. In the second stage, polling Stations are sampled from each of these AC's, again proportional to electorate size. In the final stage respondents are selected from the Electoral Rolls provided by the Election Commission. Respondents are sampled by the Systematic Random Sampling (SRS) method, which is based on a fixed interval ratio between two respondents in the polling booth. More information on the sampling and questionnaire modules of the 2004 NES can be found in Lokniti (2004). Further details about the data and sources are provided in Appendix A.4.

Tables 1 & 2 provide summary statistics for all the variables used, separately for each phase as well as over the entire sample. The summary statistics for vote shares includes information from even states that did not experience a variation in polling sequence. For the analysis, I am able to identify and estimate weights for only those states whose polling schedule is split into multiple phases. States where all polling is conducted on a single day do not contribute to identification and are not part of the sample. Therefore, to provide descriptive evidence on the degree to which vote shares of Phase 1 and Phase 2 constituencies diverged, I plot the density distributions of vote shares for each alliance in Figure 10. The figure shows that vote shares shifted in favour of the UPA, by taking away votes from the Others. The vote share distribution for the NDA alliance, however, remains stable between the two phases. In what follows, I explore this trend and provide causal evidence to show that voters in Phase 2 updated their priors in response to exit polls, more in favour of the UPA relative to the NDA.

6 Results

I begin this section by reporting results on estimated belief updating weights from aggregate data on voting returns. I, then, provide support for this evidence by examining voting outcomes at the individual level. The analysis is largely focused on examining outcomes amongst Phase 1 and Phase 2 constituencies, owing to the limited within-state variation in the polling sequence of later phases (Phase 3 and Phase 4)[37].

6.1 Belief Updating Weights

Table 3 presents estimates on belief updating weights for constituencies polling in Phase 2. While I do not report detailed results from the first-stage regression which is used to estimate priors for Phase 2 constituencies, I report the R-square for each separate equation at the bottom of the table. The regression specification, with state fixed effects and past election results, is able to capture a high level of variation in the constituency vote shares; the R-square being over 0.65 for all the regressions. Further, since constituencies are nested within states, I first check for within-state clustering of errors to determine the appropriate adjustment needed for computing standard errors. I calculate the Intra-Class Correlation (ICC) statistic for the residuals at the state level and fail to reject the zero clustering null, even at the 10% level. As a consequence, in the absence of state-level clustering, I report, normal, state and phase stratified, bootstrapped standard errors for all the following specifications in this subsection unless otherwise specified. I find that there is substantial asymmetry in the revision process. For the NDA, voters place 20% of their weight on the exit poll forecast. This effect is significant only at the 10% level; I fail to reject the no-updating null at the 5%. Estimates are more precise for voter revision vis-a-vis the UPA, in whose favour I find substantial belief updating. The results indicate that the no updating null is easily rejected at the 5% level. On average constituencies placed a weight of 0.3 on the exit poll forecast. Specifically, the log-odds of vote share received by the UP alliance increases by a factor if 0.36, when its projected log-odds of vote share increase by one unit. Rather than deriving marginal effects, I simulate counterfactual vote share distributions (in Section 6.3) using the estimated parameters to provide a sense of the relative shift in votes, from one alliance to the other, in response to the exit polls. The second panel in Table 3 reports results on within-alliance asymmetric updating. As outlined before voters are able to revise beliefs, both, when the forecasts exceed their prior expectations or when they fall below. Therefore I test whether voter revision differ between receiving good news or bad news about an alliance's performance. The results show that voters revise their priors upwards in response to 'good news' about the alliance's performance while ignoring 'bad news'. The fact that voters dismiss bad news and respond to good news could indicate the presence of underlying cognitive biases toward optimism compared to pessimism. There is a vast amount of literature that examines asymmetry in responses to negative versus positive information (Soroka 2006).

[37]Polling was split (within state) between Phase 1 and Phase 2 for seven states, between Phase 2 and Phase 3 for two states and between Phase 3 and Phase 4 for two states.

For instance Eil and Rao (2011) in a laboratory experiment find that subjects receiving negative feedback reacted very little to new signals, were far less predictable in their updating behavior and exhibited an aversion to new information whereas their inference conformed more closely to Bayes' Rule in response to good news.

Table 4 incorporates national signals and tests for whether Phase 2 voters split their updating weights between within-state forecasts (state signals) and across-state forecasts (national signals). National signals will contain more noise than state signals because voters do not observe preferences of states other than their own. For both alliances, the updating weight on the national signal is not different from zero (at the 5% level). Voters of states going to poll in more than one phase, choose to update entirely over state-level forecasts ignoring signals that they receive from other states. The extent of updating is similar to that reported in Table 3 discussed previously.

Next I report estimated belief updating weights for Phase 4 constituencies. There is some evidence for asymmetric updating with voters placing weights of 0.35 and 0.36 on exit poll forecasts respectively, for the NDA and UPA (Table 5). As before national signals are ignored. However, the weights on state-level signals are not different from zero suggesting perhaps that learning fades out during the last phases of the election with priors gaining sufficient strength. These results should nevertheless be interpreted with caution owing to the extremely small sample size (30 observations). The precision of these estimates falls rapidly when introducing further degrees of freedom. As such, the setting does not offer strong identification conditions for detecting learning in these later phases. This is an important limitation of my empirical analysis since I am unable to fully explore how observational learning evolves over time.

6.1.1 Why is Updating Asymmetric? A Simple ANOVA Test

Before moving on to present other results, I briefly discuss the interpretation of asymmetric updating result. Why did voters update much more in favour of the UPA as compared to the NDA in the 2004 elections? Intuitively, the results could be explained based on alliance manifestos and incumbency position. The incumbent alliance, NDA, endorsed a conservative center-right agenda compared to the UPA's relative liberal center-left policy position (Suri 2004; Spary and Wyatt 2006). There is a large literature in political science that discusses the effect of partisanship and incumbency on voter uncertainty and generally finds that (voters who support) liberals tend to be more ambivalent than conservatives (Feldman and Zaller 1992; Tetlock 2005). Alvarez (1998) makes two observations in this regard. Firstly, he notes that voters who are relatively more certain of the party's position are less likely to use information in their decision-making process. Traditionally, extreme right and Centre-right parties have had a support base amongst voters who tend to be relatively more certain about party's position and, therefore, less certain. More importantly, he also observes that the party challenging an incumbent always has greater uncertainty amongst voters. In the Indian context the alliance/party challenging an incumbent was the UPA.

The key aspect is that the weight on the prior for the NDA ($\beta_{NDA,t}$) will be greater than the weight on the prior for the UPA ($\beta_{UPA,t}$), if the combined variance of the noise in the prior[38] for the UPA is greater ($\sigma^2_{NDA,t} < \sigma^2_{UPA,t}$). This is evident from Eq. (41):

[38] Recall that the constituency specific prior has a combined variance given by a part of which is constant across constituencies and the other which is constant across constituencies but differs by alliance.

$$\beta_{NDA,t} = \frac{\sigma^2_{\bar{\omega}^*,t-1}}{\sigma^2_{NDA,t} + \sigma^2_{\bar{\omega}^*,t-1}} > \beta_{UPA,t} = \frac{\sigma^2_{\bar{\omega}^*,t-1}}{\sigma^2_{UPA,t} + \sigma^2_{\bar{\omega}^*,t-1}} \tag{41}$$

I use a simple ANOVA test to see whether this holds in the data. For this, I use the individual level voter post-poll data since it allows me to identify the parameters on PC fixed effects, with which I conduct the ANOVA test. Recall, that this data was collected for a sample of voters in a sample of PC's; the variance estimates are therefore sample estimates. Restricting the analysis to Phase 1 constituencies, I estimate a multinomial logit model of the voter's alliance choice with PC fixed effects. I then compute the within-state variance of the estimated PC fixed effect for each alliance. These within-state variances of the estimated PC fixed effects provide an approximate measure of $\sigma^2_{NDA,t}, \sigma^2_{UPA,t}$[39]. Table 4 reports the results on the estimate prior variance. For Phase 2 constituencies, the combined variance of the noise in the prior for NDA is 1.5 which is significantly (at the 5% level) smaller than the combined variance of the noise in the prior for the UPA which is 1.7.

6.1.2 Learning and Conformity

One issue with the interpretation of the results is whether the estimated weights reflect observational learning by voters or merely their preferences for conformity. Distinguishing between the two factors is crucial because in the absence of learning, voters will simply herd on collective actions and ignore their private signals. Parties with an early projected lead would therefore gain an early, arguably unfair, advantage. To demonstrate that this is not the case and that observation learning is indeed the mechanism by which voters update I make use of an additional source of variation in the 2004 elections.

In 2004, along with the national elections, five states also held their legislative assembly (state) elections, polling in accordance with the national election schedule. Voters residing in these states had two votes - one for the national election and one for the state elections. Parties that contested state elections included the major national parties but encompassed a wider set of state-wise smaller regional parties. By exploiting variation from states holding simultaneous legislative and general elections, I test whether voters respond to information that is orthogonal to party quality in different elections. I use the exit poll forecast of each party's performance at the national level and see whether voters update in response to such news when they vote for parties contesting state-level or assembly elections[40]. Kumar (2009) examines how voting decisions in general election and assembly election might influence each other when they are held together in India. He finds a high positive correlation between these two types of voting choices amongst voters[41]. If voters fail to *update* their assembly election voting choices, even to a small extent, despite this empirical regularity (positive correlation), then it would provide some indication that the conformity channel is absent.

Table 6 reports results from this test. The unit of observation here is an Assembly Constituency (AC). For these specifications, standard errors are adjusted for clustering within PC, since the ICC of residuals at the PC level are significant and different from zero. I, therefore, scale the standard errors by the Moulton

[39]This method is exactly equivalent to estimating Eq. (6) with state fixed effects and computing the variance of the PC level effect.

[40]Exit poll predictions of state elections are rare and almost never reported.

[41]The author also finds that there appears to be a decline in vote share of parties in assembly elections compared to general elections, potentially attributable to the presence of a larger choice set (of parties) in assembly elections.

factor (Angrist and Pischke 2009) and report these in parentheses[42]. The table shows that voters did not update their state-legislative assembly voting choices based on exit poll forecasts of the general election. This indicates that voters differ on the dimensions of uncertainty between the two types of elections and are able to distinguish their learning process based on this. If voters received direct utility from conforming with the choices of other voters within their own state then the updating weights on exit poll forecasts of the general election would have been slightly above zero reflecting this. Even if voters voted on the same dimensions between state and national elections one would have expected to see positive and significant updating[43]. Instead, despite the high correlation in voting choices for legislative assembly and general elections, this updating weight is close to zero providing some support for the observational learning hypothesis.

6.1.3 Voter Turnout

Finally, I address the important issue of voter turnout and possible endogenous assignment of constituencies into polling phases. On average 75% of the electorate voted in the 2004 election. Turnout varied by state with some states registering a turnout as low as 33%. A common concern is whether exit poll forecasts influence the probability of *voting* itself. Voters may choose not to cast their vote if they receive news that their preferred party has gained substantial leads. Conversely, unwilling voters may decide to vote and bolster the winning chances of parties projected to lose. The model and empirical results express individual choice probabilities, and therefore aggregate voting returns, to be *conditional* on the level of turnout. In this section I relax this assumption and jointly estimate voter turnout and voting returns. The objective is to take into account the inter-relationship between the decision to cast a vote and the decision to vote for a particular alliance. Therefore, the choice of alliance is *nested* under the voting decision. I specify a nested random utility model (Berry 1994) where each voter either selects the alliance who gives her the highest utility, or decides not to vote. Appendix A.3 derives the model in detail. The modified equation to estimate the belief updating weights is given by:

$$\widetilde{\vartheta}_{acs,t|\widetilde{\omega}^*_{as,t-1}} = \widetilde{\beta}_{a,t}[\widetilde{\vartheta}_{acs,t} - \delta \ln(\widetilde{\upsilon}_{acs,t-1|V})] + (1-\widetilde{\beta}_{a,t})[(1-\delta)\widetilde{\omega}^*_{as,t-1}] + \delta \ln(\widetilde{\upsilon}_{acs,t|V}) \tag{42}$$

where the $\widetilde{\vartheta}_{acs,t|\widetilde{\omega}^*_{as,t-1}}$ is the log odds of vote share received by alliance a over the entire electorate (including those who do not vote) relative to the base category, the decision to not vote. $\ln(\widetilde{\upsilon}_{acs,t-1|V})$ is the log of the vote share received by alliance a amongst voters (conditional vote shares). δ is the nesting parameter which lies between 0 and 1, and measures the correlation of the consumers' utility across alternatives belonging to the same group. It represents the extent to which alliances are considered substitutes. If $\delta = 1$, all alliances are perfect substitutes, and any shock to the system, say exit poll forecast, only determines the choice between voting and nonvoting. If $\delta = 0$, then, voters are equally likely to switch between voting for any alliance and not voting in response to an exit-poll forecast. In this specification, $(1-\widetilde{\beta}_{a,t})$ is directly comparable to the elasticity of the *conditional* vote shares with respect to the exit poll signals[44].

[42] The average number of clusters is greater than forty; the Moulton factor correction in this case performs well and can be interpreted using the t-distribution with the standard residual degrees of freedom.

[43] A finding of positive and significant updating could indicate both learning and conformity effects if voters voted on the same dimension making it impossible to distinguish between the two effects.

[44] The elasticity of conditional vote shares in the nested logit model is retrieved by scaling each coefficient by a factor of $1-\delta$. Since the transformed seat-share forecast is itself multiplied by $1-\delta$, the elasticity of the conditional vote shares with respect to seat shares is $\widetilde{\beta}_{a,t}$ whereas the elasticity of the unconditional vote shares with respect to seat shares is $(1-\widetilde{\beta}_{a,t}) \times (1-\delta)$.

Table 7 reports results from including voter turnout. The nesting parameter, δ, is estimated to be 0.577. The belief updating weight placed on the exit-poll forecast, *conditional on voting*, is estimated to be 0.23 for the UPA and this effect is significant at the 10% level. No statistically significant updating is detected for the NDA[45]. Using these results, I can calculate the belief updating weight amongst the entire electorate by multiplying each coefficient by the a factor of $1 - \delta = 0.4$. Taking into account non-voters, the results indicate that the *electorate* on average placed a belief updating weight of approximately 0.1 on exit poll forecast of the UPA and that the exit-polls results on UPA performance increased voter turnout[46].

6.1.4 Sequence Selection and Media Coverage

I also build into this estimation, a selection stage that determines the assignment probability of each constituency into polling phases. To identify the parameters for each of these equations separately, I use as exclusion restrictions, meteorological conditions that induced plausible variation in polling phase selection. There is justification for using weather related variables and in particular daily rainfall. As outlined in Section 2, the ECI in deciding polling sequence, considered the possibility of rainfall or monsoon affecting polling. They report to have consulted the India Meteorological Department for this purpose. I use this information and compile data on daily rainfall for the entire period of the election in 2004. I construct the following variables and describe how these are used to jointly estimate selection and turnout. The first variable measures the amount of average rainfall in every constituency over a +4/-4 daily window around each Phase's polling dates. I, then, use a conditional (fixed effects) logit to predict the polling phase selection probabilities for each constituency using average rainfall as a predictor. Note that the fixed effects imply that all other variables used in the system, to predict turnout, priors and coefficients on belief updating weights, are absorbed. I, therefore, do not exclude any explanatory variables used in the structural equation. Using these selection probabilities, I derive the inverse Mills ratio and include it in all equations that follow, conditioning on actual rainfall on polling day at every stage.

Table 8 & 9 report results from the procedure described above. The results, for both conditional and unconditional vote shares, are robust to selection correction and voter turnout inclusion. The selection correction terms, the inverse Mills ratio, are negative and statistically significant indicating that selection related unobservables and unobservables determining vote shares are negatively correlated. The exclusion variables used for identifying the selection equation are both highly significant. Log of rainfall in the +4/-4 window around the polling day reduces the likelihood of being selected for that polling day by 14%.

Finally, I explore whether the updating weights are sensitive to the intensity of media coverage. Using data on town-wise newspaper coverage and state-wise television coverage, I construct weights,[47] that reflect the level of media coverage in each constituency to estimate a weighted regression. Weights are constructed such that constituencies with higher media coverage are given more weight in order to analyze whether the signal's effect is greater in such areas. Table 10 reports the results from this estimation. I find that the weight placed on exit polls increases, but only slightly, when accounting for media coverage. This is expected, since the probability that voters receive exit poll results is proportional to the extent of media

[45]For both voter turnout and sequence selection results, I use the percentile confidence intervals since it is invariant to transformations of the estimated parameter. Both sets of results involve taking monotonic transformations of a subset of the parameters.

[46]The marginal effect of the exit polls on voter turnout is presented in Section 6.2.3.

[47]Appendix A.4 describes the data and procedure for constructing media weights.

coverage in their constituency. The results are exploratory since the across constituency variation in media coverage is small owing to the aggregation of town-wise media statistics at the constituency level.

6.2 Individual Voting Choice

In this section I report results from testing whether individual level voting choices are affected by surprises in alliance performance as indicated by the exit poll forecasts. Section 4.2 lays out the identification conditions required for estimating the effect. To recap, the idea is to estimate whether late voters increased their probability of voting for parties that were projected to have made 'gains', relative to early voters who received no such information[48]. To do so, I use a regression discontinuity approach and compare voting choices of individuals located on either side of the boundary that was drawn by the ECI to demarcate polling phases across constituencies. Before proceeding, I report the mean and standard deviation of the control variables in Table 2 for both Phase 1 and Phase 2 voters. A simple t-test for a difference in means shows that all variables are statistically identical along the phase boundary, except for $\#rooms$ which has a statistically significant difference, but only at the 100 km BW[49].

6.2.1 Regression Discontinuously Results

I first test for the continuity of the assignment variable, distance to the phase boundary, at the threshold. I plot both the histogram of distance to the boundary (Figure 5) and a test for density smoothness proposed by McCrary (2008) (Figure 6). The distribution of the assignment variable appears to be fairly continuous and is not discontinuous around the threshold level. I also plot local polynomial estimates of the probability of voting for the UPA and NDA in their gain states around the threshold distance. Both figures (7 & 8) suggest that Phase 2 voters increased their probability of voting for each of the alliance upon hearing news of a positive surprise.

Table 11 & 12 show regression discontinuity estimates of the treatment effect of being assigned to polling Phase 2 on the probability of voting for the UPA or NDA. For each alliance, I define 'gain' states as those states where exit polls predicted substantial improvement in alliance performance, in terms of seat share, from the previous election. For the UPA, these were the states of Andhra-Pradesh, Maharshtra and Bihar. The NDA, on the other hand, were reported to have gained in the states of Karanataka and Orissa. Voter polling in later rounds are effectively 'treatment' units since they received exit poll forecasts on such gains/losses compared to early voters who did not.

Table 11 reports the treatment effects at three different bandwidth levels, 100, 150 and 200 Kilometers. The average perimeter of an AC, the primary sampling unit for individuals in the sample, is approximately 900 Km. The choice of the bandwidths reflect this. In the absence of detailed GPS data on each individual's location, I choose bandwidths that contain an adequate number of AC clusters for analysis. In every panel of Table 11 & 12 I report the number of sample AC clusters together with the sample size at each bandwidth level. The upper panel reports results from a specification which includes polynomials of latitude and longitude, while the bottom panel uses distance to the boundary polynomials as a smoothing

[48] I focus on gains instead of losses as aggregate results reported in Section 6.1 suggest that voters react to good news and dismiss exit poll results on negative performance.

[49] As before I focus on the difference in outcomes between Phase 1 and Phase 2 voters. I limit the analysis to these phases because no major gains were reported for any alliance amongst states that split their polling between Phase 2-3 and Phase 3-4.

function. All specifications control for boundary-segment fixed effects. Results in columns 4, 5 and 6 also include interactions of the distance/latitude-longitude variables with the boundary-segment fixed effects. Finally, all specifications adjust standard errors for clustering at the Assembly Constituency and boundary segment level and report marginal effects from a logit specification. The average treatment effect is fairly stable though most of the specifications, its estimates ranging between 0.2 and 0.3. Voters in Phase 2 constituencies, who receive forecasts about UPA gains, increase their probability of voting for the UPA by approximately 25%. In contrast, consistent with the results from aggregate voting returns, voters place no weight on receiving news about NDA gains. Table 12 shows that the effect of receiving news on NDA gains on the probability of voting for the NDA is close to zero.

As a robustness check, I also provide results from reducing the bandwidths to 50 and 75 Kilometres respectively in Table 13. The results are robust to narrowing band-widths and follow the same pattern of results previously reported. Voters in Phase 2 increase their probability of voting for UPA in UP gain states but fail to update for NDA in NDA gain states. In fact the magnitude of the treatment effect increases slightly at the lowest bandwidth of 50 Km.

6.2.2 Learning, Conformity and Placebo Check

Another robustness check that I carry out in order to rule out conformity effects, is to test whether voter's update their assembly election voting choice in response to predicted alliance gains at the national level. This is similar to the test employed with aggregate voting returns data where I showed that voters did not update their state-legislative assembly voting choices based on exit poll forecasts of the general election. Columns 1, 2 and 3 of Table 14 validate this finding at the individual level. The results show that voters did not update their probability of voting for the UPA in the assembly elections in UPA gain states. This result is robust to using polynomials of latitude/longitude and distance as well as to the choice of bandwidth. Note that the regression results also indicate, to some extent, that there were no state-specific party shocks/trends between the two polling phases that were set apart by a gap of one week. Such a trend could have caused shifts in the vote share distributions, of both general and assembly elections, independent of exit poll results. Given that the coefficient on the treatment effect is not different from zero, I rule out this possibility.

An additional placebo check that I employ is to estimate regression discontinuity based treatment effects for voting choice in the 1999 General Election. I use the recall question on 1999 voting choices from the NES survey to conduct this check. The assumption I make in the regression discontinuity design is that treatment and control units (Phase 2 and Phase 1 constituencies) differ only in their ex-post outcomes (voting choices), conditional on the variables determining their selection into treatment. This means that conditional on the 1999 Phase allocation, the 2004 Phase allocation should not have any effect on ex-ante voting choice made in 1999. If this is so, then it is reasonable to believe that the treatment and control units differ on many other ex-ante outcomes, making the identification strategy invalid. Columns 3, 4 and 5 of Table 14 show that this is not the case. The treatment effect estimates are mostly statistically insignificant indicating that prior voting choice did not differ between Phase 1 and Phase 2 constituencies.

6.2.3 Voter Turnout

I now present regression discontinuity results[50] that incorporate the decision to vote. I estimate a nested logit model (structure is described in Appendix A.3) of voting choice. In the first stage the individual chooses between voting and not voting; conditional on her decision to vote, she chooses between voting for each of the alliances. Table 15 reports results from the nested logit model. Recall that a low value of the nesting coefficient, δ, means that voters consider alliances to be dissimilar (for unobserved reasons). This coefficient is estimated to be 0.4 suggesting that there is some, but not complete, amount of substitutability between alliance choices. The estimate is close, but not equal, to the nesting coefficient estimated from the aggregate data[51] (0.57).

The upper panel of Table 15 shows that individuals increased their probability of voting for the UPA in Phase 2 relative to Phase 1 in states where the UPA was expected to make substantial gains. The coefficient measures the log-odds of voting for each alliance relative to the base category of non-voting. To facilitate interpretation, I use the parameters of the model and report the predicted change in the probability of voting for an alliance when the treatment variable changes from zero to one, holding other variables at their mean values. The coefficient on the treatment effect suggests two things. Firstly, (conditional on the decision to vote) *voters* increased, significantly, their probability of voting for the UPA by 16%. Secondly, individuals increased their probability of voting by approximately 25% upon receiving news of a UPA gain. The bottom panel of the same table reports voting choices in NDA gain states. As before, I find no significant treatment effects, i.e., the effect of receiving news on NDA gains on the probability of voting for the NDA and the probability of voting itself, is close to zero.

While the above results provide some indication that exit polls have an effect on voter turnout, they do not confirm the mechanism of voter learning. For instance, one could imagine that exit polls could have a direct effect on voter turnout, not by signalling party quality, but by influencing voters' perceptions about the *closeness* of the election. Sudman (1986) points out that indeed, exit polls may encourage some some individuals to vote when the election (prediction) is not close because they are sure to be on the winning side (a bandwagon effect) while others may vote to reduce the magnitude of the anticipated winning margin (an underdog effect). However, the multi-phase procedure of the Indian elections together with the long duration over which they are held ensure to some extent that this is not the case and that voters are unable to full perceive the closeness of the overall election, at least in the early phases. For instance, the first phase of election saw a total of 141 constituencies going to poll. This meant that Phase 2 voters received information on the voting patterns of 26% of the overall polling constituencies. Further, since each constituency represents a single seat in the parliament, voters essentially received information on only 26% of the contested seats in parliament. In the absence of information on the remaining 74% seats, voters were unable to gauge perfectly the apparent closeness of the overall election.

Nevertheless, to provide further clarity on this issue, I examine the interaction effects of exit poll predictions and individual preferences vis-a-vis an individual's decision to vote. I use each individual's voting choice in the previous election (1999) as a proxy for her ideological preference. I then test whether exit polls had a differential effect on each individual's decision to vote given her own ideological preference.

[50] Since I use an RDD design to tackle the endogenous assignment into treatment groups, I do not undertake a further sequence selection correction as done for the aggregate voting returns case.

[51] The nesting parameter is estimated for the full sample of Phase 1 and Phase 2 voters.

I undertake this test for the full sample of Phase 1 and Phase 2 voters and under three different forecast conditions: when exit polls predict reported gains for an alliance ('Surprise'), when exit polls predict that an alliance is expected to win more than half the seats in the voters' state ('Lead') and when exit polls predict that an alliance is expected to win more than one-third the seats in the voters' state ('Narrow Win'). Table 7 reports results from this exercise. I find that, as before, exit-poll predictions on a UPA surprise increase the probability of voting. Conversely an NDA surprise has no effect on the probability of voting. More importantly I find that none of the interaction effects are significant. The prediction of a surprise has no differential effect on individuals who voted either UPA or NDA in the previous election (voting for 'Others' being the base category). I also find limited effects of predicted 'Leads' or 'Narrow Wins' on the probability of voting, confirming the main predictions of the paper – voters respond mainly to surprises in voting returns.

6.2.4 Geographical Discontinuously Results

Finally, I report results obtained from estimating treatment effects at *each point* of the polling schedule boundary as implied by Eq. (40). I first plot the distribution of coefficient estimates (treatment effects), corresponding to each boundary point, in Figure 12 against its q-values (corrected critical p-values). The figure shows the discovery set, which is defined as the subset of p-values that are at or below the corrected critical p-value (Newson 2010), under *any* FDR. There is one panel for each bandwidth used. The horizontal axis gives the coefficient on the treatment effect and the vertical axis gives the corresponding q-value on a reverse log scale. The horizontal reference line (in dashed black line) indicates the 95% confidence level. All points above this line are statistically significant at 5%. The figure shows that, in UPA gain states, there are a large number of statistically significant treatment effect coefficients and most of these occur at a bandwidth greater than 150. This is not surprising given the low precision power of the 150 Km bandwidth whose sample contains only 12 AC clusters on average. In contrast very few treatment effects for NDA states are statistically significant, which is in line with earlier results,

In Table 17, I calculate the average treatment effect based on the distribution of point estimates, at each bandwidth. It can be seen that for bandwidths greater than 200, the average treatment effects, measuring the marginal increase in the probability of voting for UPA in UPA gain states amongst later voters, range from around 0.18 to 0.25. The effect is almost zero, or even negative at certain bandwidths, for the NDA. Largely, the results are consistent with those obtained using the semi-parametric regression discontinuity approach. To finish off, I plot each treatment effect against its boundary point on the map given in Figure 9. All significant (at the 5% level) treatment effect estimates are coloured in red. The figure is useful for visualizing the extent of within-state heterogeneity in treatment effects.

6.3 Simulated Vote Shares

I end the results section by presenting simulated vote shares using the estimated belief updating weights. The objective is to assess the direction and magnitude of a shift in vote shares of each alliance due to the release of exit-poll results. In order to obtain predictions, I estimate a multinomial logit model of voting choice for UPA gain states, using the individual NES post-poll data (Eq. (35)) where the dependent variable is multinomial rather than binary). The variable of interest, the treatment effect, is a positive surprise in the form of expected gains to the UPA alliance. To simulate the counterfactual vote for each person in the sample, I set the treatment effect variable to zero (no exit polls) letting other characteristics of the voter take on their true values, and use the predicted value algorithm (King et al. 2000)[52]. For comparison, I also simulate votes when exit polls are allowed and the treatment effect takes the mean value, as observed in the sample. Using this procedure, I generate 1000 simulated election outcomes.

Figure 11 displays the results of the simulations. The figure is a "ternary plot" (Katz and King 1999) whose coordinates represent predicted fractions of the vote received by each alliance under different simulated election outcomes. For example, the points coloured in black show the density distribution of simulated vote shares when voters receive exit poll results. The mean predicted vote proportion in this scenario is approximately 0.35 (NDA), 0.46 (UPA) and 0.16 (Others). The counterfactual occurs when there are prohibitions on the release of exit polls and voters do not receive any news. In this case, the vote share distribution shifts towards the left (in gray) and the mean predicted vote proportion is approximately 0.4 (NDA), 0.31 (UPA) and 0.27 (Others). This means that the positive surprise news effect for the UPA, increased UPA vote share by shifting votes away from 'Others'. The vote shares of the NDA remains fairly stable, reducing only slightly in response to this news. This is consistent with the descriptive evidence (Figure 10) and the overall results that indicate little to no updating for the NDA.

7 Conclusion

In this paper I have examined whether the mid-election release of exit poll results affects voting behaviour. I use a social learning framework to explain how voters, who are imperfectly informed and have uncertainty about party quality, can use exit poll predictions of electoral outcomes as noisy signals to update over their prior in a Bayesian manner. I test the social learning model as well as some aspects of observational learning, empirically, by exploiting the multi-phase polling schedule of the Indian general elections. I find two main results: voters polling in later phases do react to exit poll results, in a limited and asymmetric way. Voters only update their priors in response to predictions based on early round voters from within their own states and dismiss predictions about across-state voting patterns. Further the extent to which voters update varies across parties. The latter effect depends largely on the strength of the priors that voters hold for each party. As a consequence, positive surprises in party performance, in the form of expected gains and losses, shift the vote share distribution substantially. I confirm these results both using aggregate voting returns data and individual post-poll survey data.

While the paper attempts to address the specific policy question on whether and how *exit polls* can influence voting choice, its implications are applicable to many settings where media platforms can alleviate the uncertainty of imperfectly informed agents. Finally, future research can be extended in several directions. It will be interesting to examine the interaction effects of opinion polls and exit polls in resolving voter uncertainty. So far in the paper, I have assumed that all voters receive only private signals that are used to form initial priors. However opinion polls can be modelled to act as pre-election public signals such that they contribute to shaping a common prior. Subsequently, one can analyse the divergence of public opinion as mid-election exit-poll results are released. Another related question is how mid-election news and media coverage about politician and party specific scandals affect electoral outcomes. In particular, it would be worth exploring whether voters react, if at all, to specific types of news (personal

[52]I use the CLARIFY package provided by Tomz et al. (2003).

scandals or in-office scandals) varied by the intensity of media coverage. Detailed data on news events and television/newspaper coverage over the election period can be used to potentially answer the latter question.

References

ABJORENSEN, N. (2012): "How polls are shaping the political process," *in Political Polling In Asia-Pacific (eds. Carthew, A and Winkelmann, S)*, 31–39.

ACEMOGLU, D. AND A. OZDAGLAR (2011): "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, 1, 3–49.

AIDT, T., M. GOLDEN, AND D. TIWARI (2011): "Incumbents and Criminals in the Indian National Legislature," *Cambridge Working Papers in Economics*.

ALI, S. AND N. KARTIK (2011): "Herding with collective preferences," *Economic Theory*, 1–26.

ALVAREZ, R. (1998): *Information and elections*, Univ of Michigan Pr.

ALVAREZ, R. AND J. NAGLER (1998): "When politics and models collide: Estimating models of multiparty elections," *American Journal of Political Science*, 55–96.

ANAND, S. AND C. JENKINS (2004): "Exit Polls: Debating Freedom or Fairness," *Economic and Political Weekly*, 4971–4973.

ANGRIST, J. AND J. PISCHKE (2009): *Mostly harmless econometrics: An empiricist's companion*, Princeton Univ Pr.

BANDIERA, O. AND I. RASUL (2006): "Social networks and technology adoption in northern mozambique," *The Economic Journal*, 116, 869–902.

BENJAMINI, Y. AND Y. HOCHBERG (1995): "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.

BENJAMINI, Y. AND D. YEKUTIELI (2001): "The control of the false discovery rate in multiple testing under dependency," *Annals of statistics*, 1165–1188.

BENNETT, D., C. CHIANG, AND A. MALANI (2011): "Learning During a Crisis: the SARS Epidemic in Taiwan," *NBER Working Papers*.

BERRY, S. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 242–262.

BLACK, S. (1999): "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 114, 577–599.

BYRNE, G., M. CHARLTON, AND S. FOTHERINGHAM (2009): "Multiple dependent hypothesis tests in geographically weighted regression," in *10th International Conference on GeoComputation, UNSW, Sydney, November-December*.

CAI, H., Y. CHEN, AND H. FANG (2009): "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, 99, 864–882.

CARDELL, N. (1997): "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 13, 185–213.

CHIANG, C. AND B. KNIGHT (2011): "Media bias and influence: Evidence from newspaper endorsements," *The Review of Economic Studies*, 78, 795–820.

CRIPPS, M., J. ELY, G. MAILATH, AND L. SAMUELSON (2008): "Common Learning," *Econometrica*, 76, 909–933.

DELL, M. (2010): "The persistent effects of Peru's mining Mita," *Econometrica*, 78, 1863–1903.

DEMICHELIS, S. AND A. DHILLON (2010): "Learning in elections and voter turnout," *Journal of Public Economic Theory*, 12, 871–896.

ECI (2004): "General Elections - 2004 Schedule Of Elections," *Press Note*.

——— (2009): "General Elections - 2009 Schedule Of Elections," *Press Note*.

EIL, D. AND J. RAO (2011): "The good news-bad news effect: Asymmetric processing of objective information about yourself," *American Economic Journal: Microeconomics*, 3, 114–138.

FELDMAN, S. AND J. ZALLER (1992): "The political culture of ambivalence: Ideological responses to the welfare state," *American Journal of Political Science*, 268–307.

FOTHERINGHAM, A., C. BRUNSDON, AND M. CHARLTON (2002): *Geographically weighted regression: the analysis of spatially varying relationships*, John Wiley & Sons Inc.

GELMAN, A. AND G. KING (1994): "A unified method of evaluating electoral systems and redistricting plans," *American Journal of Political Science*, 514–554.

GENTZKOW, M. (2006): "Television and voter turnout," *The Quarterly Journal of Economics*, 121, 931–972.

GENTZKOW, M., J. SHAPIRO, AND M. SINKINSON (2011): "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 101, 2980–3018.

GOI (2009): "The Representation Of The People (Amendment) Act," *Ministry of Law and Justice*.

GREEN, D., A. GERBER, AND S. DE BOEF (1999): "Tracking opinion over time: A method for reducing sampling error," *Public Opinion Quarterly*, 178–192.

GUL, F. (1998): "A Comment on Aumann's Bayesian view," *Econometrica*, 66, 923–927.

IMBENS, G. AND T. LEMIEUX (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

KARANDIKAR, R., C. PAYNE, AND Y. YADAV (2002): "Predicting the 1998 Indian parliamentary election," *Electoral Studies*, 21, 69–89.

KATZ, J. AND G. KING (1999): "A statistical model for multiparty electoral data," *American Political Science Review*, 15–32.

KEELE, L. AND R. TITIUNIK (2011): "Geographic Boundaries as Regression Discontinuities," *mimeo*.

KING, G. (1990): "Electoral responsiveness and partisan bias in multiparty democracies," *Legislative Studies Quarterly*, 159–181.

KING, G., M. TOMZ, AND J. WITTENBERG (2000): "Making the most of statistical analyses: Improving interpretation and presentation," *American Journal of Political Science*, 347–361.

KNIGHT, B. AND N. SCHIFF (2010): "Momentum and Social Learning in Presidential Primaries," *Journal of political economy*, 118, 1110–1150.

KUMAR, S. (2009): "Do General and State Election Results Influence Each Other?" *Eternal India*, 1.

LAHIRI, A. AND P. ROY (1984): "Assessing swings in multi-party systems: The Indian experience," *Electoral Studies*, 3, 171–189.

LARSON, N. (2011): "Inertia in social learning from a summary statistic," *MPRA Paper*.

LEE, D. AND E. MORETTI (2009): "Bayesian learning and the pricing of new information: Evidence from prediction markets," *The American Economic Review*, 99, 330–336.

LINZER, D. (2012): "The Relationship between Seats and Votes in Multiparty Systems," *forthcoming in Political Analysis*.

LOKNITI, T. (2004): "National Election Study 2004: An Introduction," *Economic and Political Weekly*, 5373–5382.

LYBBERT, T., C. BARRETT, J. MCPEAK, AND W. LUSENO (2007): "Bayesian herders: Updating of rainfall beliefs in response to external forecasts," *World Development*, 35, 480–497.

MCALLISTER, I. AND D. STUDLAR (1991): "Bandwagon, underdog, or projection? Opinion polls and electoral choice in Britain, 1979–1987," *The Journal of Politics*, 53, 720–741.

MCCRARY, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

MORETTI, E. (2011): "Social learning and peer effects in consumption: Evidence from movie sales," *The Review of Economic Studies*, 78, 356.

MUNSHI, K. (2004): "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution," *Journal of Development Economics*, 73, 185–213.

NEVO, A. (2000): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513–548.

NEWSON, R. (2010): "Frequentist q-values for multiple-test procedures," *Stata Journal*, 10, 568.

NIEMI, R. AND P. FETT (1986): "The Swing Ratio: An Explanation and an Assessment," *Legislative Studies Quarterly*, 75–90.

RAJEEVAN, M., J. BHATE, J. KALE, AND B. LAL (2006): "High resolution daily gridded rainfall data for the Indian region: Analysis of break and active monsoon spells," *Current Science*, 91, 296–306.

RANGARAJAN, M. AND P. VIJAI (2002): "INDIA: First Past the Post on a Grand Scale," *The International IDEA Handbook of Electoral System Design (Hrsg.: Reynolds, Andrew and Reilly, Ben), Schweden*, 32–35.

SOROKA, S. (2006): "Good news and bad news: Asymmetric responses to economic information," *Journal of Politics*, 68, 372–385.

SPARY, C. AND A. WYATT (2006): "The general election in India, May 2004," *Electoral Studies*, 25.

SRIDHARAN, E. (2004): “Electoral coalitions in 2004 general elections: Theory and evidence,” *Economic and Political Weekly*, 5418–5425.

SUDMAN, S. (1986): “Do exit polls influence voting behavior?” *Public Opinion Quarterly*, 50, 331–339.

SURI, K. (2004): “Democracy, economic reforms and election results in India,” *Economic and Political Weekly*, 5404–5411.

TETLOCK, P. (2005): *Expert political judgment: How good is it? How can we know?*, Princeton Univ Pr.

THEIL, H. (1969): “The desired political entropy,” *The American Political Science Review*, 63, 521–525.

THIESSEN, A. (1911): “Precipitation averages for large areas,” *Monthly weather review*, 39, 1082–1089.

TOMZ, M., J. TUCKER, AND J. WITTENBERG (2002): “An easy and accurate regression model for multiparty electoral data,” *Political Analysis*, 10, 66–83.

TOMZ, M., J. WITTENBERG, AND G. KING (2003): “CLARIFY: Software for interpreting and presenting statistical results,” *Journal of Statistical Software*, 8, 1–30.

ZELLNER, A. (1962): “An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias,” *Journal of the American statistical Association*, 348–368.

A Appendix

A.1 Exit Polls - Forecasting Techniques and Regulation

In this section, I discuss two important aspects of the exit polls. I first describe the forecasting methodology used by many polling agencies and then discuss the current policy debate on its regulation.

In terms of technique, most polling agencies aim to predict the number of seats won by each party in the Parliament. As an example, in Figure 2, I reproduce a newspaper clipping from the Times of India, a nationally circulated newspaper, reporting exit poll results immediately after Phase 1 of the elections. Forecasting election outcomes based on sample data is not an easy task; 'pollsters' face tremendous logistic and methodological challenges in attempting to produce these estimates. Without intending to understate the complexities of the procedures, I provide a short summary of the general forecasting and estimation technique employed by many agencies that conduct exit polls. The main methodological challenge faced by most polling agencies is of converting estimated votes into seats. Karandikar et al. (2002) provide a detailed account of methods used to predict the results of the 1998 Indian parliamentary election by the CSDS team. Using a stratified nationally representative sample of consistencies and voters, they predict the percentage votes that a major party is to get in the election. They do so by estimating the sample percentage of vote for each of the major parties over the state and then take its difference from the average vote in the previous election. Applying this change to the actual vote share in the previous election provides them with the predicted vote share for each party in each constituency. They, then, aggregate party votes by alliance and convert these estimated vote shares into seat level predictions at the state level. To do so, they employ a probabilistic method that produces probabilities that each party will win the seat. This method accounts for state specific 'swing factors' (discussed in Section 3.2) so that probabilities take into account the margin of victory between the first and second place parties. Since most agencies follow more or less the same outlined procedure, differing only in their sample draws, variation in estimates across polling agencies tends to be quite low. This means that voters, on average receive similar information regardless of their choice of agency/media-outlet from which they seek to hear exit poll predictions.

I finish off this section by briefly discussing the regulatory status of exit-poll broadcasting, as it currently stands in India. Overall, till today, any attempt to regulate the conduct of opinion polls and exit polls is challenged on the grounds that it would fundamentally violate Articles 19(1)(a & g) of the Constitution of India, namely the right to freedom of speech and expression and the right to practise any profession[53]. The ECI had attempted to prohibit the publication of exit polls and succeeded in imposing the ban for the first two phases of the 1999 elections. However, its decision was later overturned by the Supreme Court which stated that the ECI Guidelines "exceeded the power of 'superintendence, direction and control' granted to it by Article 324 of the [Indian] Constitution." (Note 13, pp. 18-19). Exit polls were immediately broadcast on national television at the end of the third phase. As mentioned before, it was only in 2008 that the Government of India approved the ECI's proposal by amending the Representation of People Act of 1951 and inserting section 126 (A & B) stating: "No person shall conduct any exit poll and publish or publicise

[53]This position was taken by the Judiciary in 2004. The Attorney General of India, at the time, expressed that he took "serious and formidable objection" to the Commission's proposal for a ban on the publication and telecast of opinion and exit polls during elections since it would "be violative of Article 19 (1) (a) of the Constitution" . He recommended, instead, that the ECI "issue directions requiring the media to comply...which would be regulatory in nature and not restrictive of the fundamental right of free speech and expression under Article 19(1)(a)."('*Ban on opinion, exit polls unconstitutional, says Soli Sorabjee*', Hindu, April 10 2004).

by means of the print or electronic media or disseminate in any other manner, whatsoever, the result of any exit poll during such period, as may be notified by the Election Commission in the regard."(GOI 2009, pp. 2). The inclusion of this amendment continues to be controversial since the time of its enactment. Anand and Jenkins (2004) had earlier argued that the phraseology of the constitutional amendment could prove more problematic than beneficial. This is because while the amendment prohibits conduct and publication of exit polls during the period starting from 48 hours before the close of poll in an election, it does not cover opinion polls. The ECI, therefore, continues to seek a regulatory ban on opinion polls as well, claiming that the results of opinion poll can be published even on the day of election polling and that there is no provision of law to restrict dissemination through print media (Notification No.3/1/2010/SDR, ECI). On the other hand, many consider the ECI's proposals as being far too radical, especially given the lack of evidence on voter influence from exit/opinion poll results.

A.2 Seat Share to Vote Share Mapping

In this section, I describe how seat shares can be mapped on to a given distribution of forecasted vote shares for each state in a multiparty/multi-alliance set up. To do this, I use a commonly used seats-votes formulation that specifies a relationship between the seat share obtained and the overall vote share of any party as a function of electoral 'responsiveness'. Electoral responsiveness is popularly termed as the 'swing factor/ratio'. It is defined as the "percentage change in legislative seats associated with a one percent change in legislative votes" (Niemi and Fett 1986, pp.76)[54]. In other words, the swing factor measures the extent to which legislative seats are allocated to parties in proportion to the division of citizens votes (King 1990).

The general form of the seats-votes relationship for two-party systems is given by:

$$\frac{S}{1-S} = \left(\frac{V}{1-V}\right)^{\rho} \tag{43}$$

where S and V are seat share and vote share respectively and ρ is the electoral responsiveness parameter. Values of $\rho > 1$ are considered highly responsive and majority parties are expected to receive a 'bonus' of seats over their vote share. A value of $\rho = 1$ demonstrates proportional representation in which seats are allocated according to vote share received. A special case of the general form occurs when the sear-votes curve follows a cube law and ρ takes the value three (Theil 1969)[55]. King (1990) derives a model to generalize this relationship and extends it to multi-party systems. I use this model and derive the seats-votes relationship for every given state at each phase of the election, to show how news about *seat share* forecasts can be used to inform voters about underlying vote shares.

Following King (1990), I denote S_{as} as the random variable for the number of seats allocated to alliance a in state s. Its observed realization is s_{as}. The total number of seats within a given state is then given by $s_s = \sum_{a=1}^{k} s_{as}$. As defined before, v_{acs} is the proportion of votes received by alliance a in constituency c of state s. Its overall vote share in the given state, v_{as} is given by averaging v_{acs} across all constituencies. Finally let π_{as} be the average probability that one of the s_s, seats will be allocated to alliance a in state s. The numbers of seats allocated to the different parties/alliances follows a multinomial probability

[54] See Linzer (2012) for an excellent discussion of swing ratios in different contexts.
[55] See Theil (1969) for an axiomatic and information theoretic approach to deriving Eq. (43).

distribution where the random choice being made is that of the electoral system in assigning seats to parties[56]. This is given by:

$$
\begin{aligned}
(S_{1s}, S_{2s}, ...S_{ks}) &\sim \text{Multinomial}(s_s|\pi_s) \\
&= s_s! \prod_{a \in A_s} \frac{\pi_{as}^{s_{as}}}{s_{as}!}
\end{aligned} \tag{44}
$$

Thus, the expected seat shares for alliance a in state s is given by[57]:

$$
\begin{aligned}
\frac{E(S_{as})}{s_s} = \pi_{as} &= \frac{\upsilon_{as}^{\rho_s}}{\sum_{k=0}^{A} \upsilon_{ks}^{\rho_s}} \\
&= \frac{\exp(\rho_s \ln \upsilon_{as})}{\sum_{k=0}^{A} \exp(\rho_s \ln \upsilon_{ks})}
\end{aligned} \tag{45}
$$

where ρ_s is the state-level swing factor that I assume to be common knowledge amongst voters within a given state. To facilitate comparison with the previously defined log-odds of vote share, I apply a simple transformation and divide the state level vote share of each party by the state level vote share of the base party i.e. $z_{as} = \frac{\upsilon_{as}}{\upsilon_{0s}}$. Eq. (45) can now be rewritten as:

$$
\pi_{as} = \frac{\exp(\rho_s \ln z_{as})}{\sum_{k=0}^{A} \exp(\rho_s \ln z_{ks})} \tag{46}
$$

Using Eq. (46) and noting that $z_{0s} = 1$, we can define the seat share probability of alliance a relative to the baseline alliance's seat share as:

$$
\begin{aligned}
\frac{\pi_{as}}{\pi_{0s}} &= \frac{\exp(\rho_s \ln z_{as})}{\exp(\rho_s \ln z_{0s})} \\
\ln \frac{\pi_{as}}{\pi_{0s}} &= \rho_s \ln z_{as} \\
&= \rho_s \ln \left(\frac{\upsilon_{as}}{\upsilon_{0s}} \right) \\
&= \rho_s \vartheta_{acs}
\end{aligned} \tag{47}
$$

Since the exit polls project an *estimate* of the true vote shares (given by Eq. (12)), I define a seat share forecast denoted by $\omega^*_{asm,1}$. The seat share forecast based on voters in Phase 1 constituencies is given by:

$$
\begin{aligned}
\omega^*_{asm,1} = \frac{1}{\rho_s} \ln \frac{\pi^*_{as,1}}{\pi^*_{0s,1}} &= \vartheta^*_{as,1} \\
\omega^*_{asm,1} &= q_a + \xi_{as} + \eta_{as} + \kappa_{sm}
\end{aligned} \tag{48}
$$

[56]King (1990) notes that the assumption of independence of irrelevant alternatives is satisfied here, since the entire stochastic component is conditional on all parties and votes.

[57]Note that I use the expression υ_{as} to denote the aggregate vote share received by alliance a at the *state* level. This is derived by aggregating the logit model in Eq. (6) over the state, rather than the constituency as previously done.

A.3 Voter Turnout

In this section, I derive the expression for belief updating weights amongst the entire electorate. To do so, I adopt the nested logit model as provided by Berry (1994). The nested logit model classifies the electorate into two groups/choice-sets, voters (V) and non-voters (NV). Each choice-set is further sub-divided into various alternatives. Denote the set of alternatives in each group as $\mathcal{J}_V$. Since the decision to not vote contains no further alternatives, it serves as a natural base category and is considered as an 'outside option'. An outside option is assumed to be the only member of its group. Alliance choices are nested within the group of voters. Therefore, each individual i will vote for alliance a conditional upon the decision to vote. To incorporate this nesting structure, I modify Eq. (49) as:

$$E_1(u_{aics,1}|\theta_{acs},\eta_{as},\psi_{aics}) = \theta_{acs} + \eta_{as} + \varsigma_{Vics} + (1-\delta)\psi_{aics} \tag{49}$$

where the variable ς is common to all alternatives within the group of voters and can be thought of as the unobserved utility from voting. It has a distribution function that depends on δ, such that if ψ is an extreme value random variable, then $[\varsigma + (1-\delta)\psi]$ is also an extreme value random variable (Cardell 1997). In the Industrial Organization literature, the nesting parameter δ lies between 0 and 1 and measures the correlation of the consumers' utility across alternatives belonging to the same group. In this context, it represents the extent to which alliances are considered substitutes. If $\delta = 1$, all alliances are perfect substitutes, and any shock to the system, say exit poll forecast, only determines the choice between voting and nonvoting. If $\delta = 0$ the voters are equally likely to switch between voting for any alliance and not voting in response to an exit-poll forecast. In this case, the standard mutinomial logit model applies, where the choice of not voting competes symmetrically with the choice of voting for a particular alliance. I estimate the value of δ from the data to determine the appropriate logit structure.

To derive the conditional and unconditional vote shares, I slightly modify notation and define the following terms:

$$\begin{aligned}
\widetilde{\upsilon}_{acs,1} &= \text{Unconditional total vote share of alliance } a \text{ amongst the electorate,}\\
\widetilde{\upsilon}_{Vcs,1} &= \text{Fraction of electorate that votes (Voters),}\\
\widetilde{\upsilon}_{acs,1|V} &= \text{Alliance share of total votes cast.}
\end{aligned}$$

Normalizing the mean utility level for the outside option (not voting) to be zero, it can be seen that $\widetilde{\upsilon}_{acs,1} = \widetilde{\upsilon}_{acs,1|V} \times \widetilde{\upsilon}_{Vcs,1}$. As before, the alliance share amongst the voting group is given by the modified Eq. (7):

$$\widetilde{\upsilon}_{acs,1|V} = \frac{\exp\left[\frac{\theta_{acs}+\eta_{as}}{(1-\delta)}\right]}{\sum_{k\in\mathcal{J}_V}\exp\left[\frac{\theta_{kcs}+\eta_{ks}}{(1-\delta)}\right]} \tag{50}$$

Denoting the denominator of the above expression as $I_{Vcs,1} = \sum_{k\in\mathcal{J}_V}\exp\left[\frac{\theta_{kcs}+\eta_{ks}}{(1-\delta)}\right]$, I obtain the probability that an individual will choose to vote (share of voters) as:

$$\widetilde{\upsilon}_{Vcs,1}(\theta_{acs},\eta_{as},\delta) = \frac{I_{Vcs,1}^{(1-\delta)}}{\sum_V I_{Vcs,1}^{(1-\delta)}} \tag{51}$$

$$\widetilde{\upsilon}_{acs,1} = \frac{\exp\left[\frac{\theta_{acs}+\eta_{as}}{1-\delta}\right]}{I^{\delta}_{Vcs,1}(\sum_V I^{1-\delta}_{Vcs,1})} \tag{52}$$

$$\widetilde{\vartheta}_{acs,1} = \ln\left(\frac{\widetilde{\upsilon}_{acs,1}}{\widetilde{\upsilon}_{0cs,1}}\right) = \theta_{acs} + \eta_{as} + \delta\ln(\widetilde{\upsilon}_{acs,1|V}) \tag{53}$$

To derive a similar expression for exit poll results, I rewrite Eq. (53) for the seat share equation corresponding to Eq. (46). Recall that the expression for expected seat shares for alliance a in state s is given by:

$$\widetilde{\pi}_{as,1} = \frac{\exp(\rho_s \ln\widetilde{z}_{as,1})}{\sum_{k=0}^{A}\exp(\rho_s \ln\widetilde{z}_{ks,1})} \tag{54}$$

where $\widetilde{z}_{as} = \frac{\widetilde{\upsilon}_{as,1}}{\widetilde{\upsilon}_{0s1}}$. Taking the logarithms of Eq. (54) and substituting with Eq. (53), we get:

$$\ln(\widetilde{\pi}_{as,1}) - \ln(\widetilde{\pi}_{0s,1}) = \rho_s \ln\left(\frac{\widetilde{\upsilon}_{as,1}}{\widetilde{\upsilon}_{0s,1}}\right) \tag{55}$$

$$\frac{1}{\rho_s}\ln\left(\frac{\widetilde{\pi}_{as,1}}{\widetilde{\pi}_{0s,1}}\right) = q_{as} + \xi_{as} + \eta_{as} + \delta\ln(\widetilde{\upsilon}_{as,1|V}) \tag{56}$$

I now derive the expression for $\widetilde{\upsilon}_{as,1|V}$ in terms of seat share forecasts. Note that the nesting structure applies only to voting behaviour. The relationship between the unconditional seat shares and the conditional seat shares is given by the simple multinomial logit model. I still retain two groups, voter and non-voters and hypothetically assume that non-voters receive a certain proportion of seats. To see this, note that the conditional seat shares $\widetilde{\pi}_{as|V}$ are allocated according to the conditional vote shares in the following way (analogous to Eq. (45)):

$$\begin{aligned}
\widetilde{\pi}_{as,1|V} &= \frac{\exp(\rho_s \ln\widetilde{\upsilon}_{as,1|V})}{\sum_{k=0}^{A}\exp(\rho_s \ln\widetilde{\upsilon}_{ks,1|V})} \\
\ln(\widetilde{\pi}_{as,1|V}) &= \rho_s \ln\widetilde{\upsilon}_{as,1|V} - \ln(\sum_{k=0}^{A}\exp(\rho_s \ln\widetilde{\upsilon}_{ks,1|V})) \\
\ln(\widetilde{\pi}_{as,1|V}) &= \rho_s \ln\widetilde{\upsilon}_{as,1|V} - \ln(R_{Vs,1})
\end{aligned} \tag{57}$$

where $R_{Vs,1}$ denotes the inclusive value. The total share seats allocated to voters and non-voters is given by:

$$\begin{aligned}
\widetilde{\pi}_{Vs,1} &= \frac{R_{Vs,1}}{\sum_V R_{Vs,1}} \\
\widetilde{\pi}_{0s,1} &= \frac{1}{\sum_V R_{Vs,1}} \\
\ln\widetilde{\pi}_{Vs,1} - \ln\widetilde{\pi}_{0s,1} &= \ln(R_{Vs,1})
\end{aligned} \tag{58}$$

Substituting Eq. (58) in Eq. (57) and noting that $\tilde{\pi}_{as|V} = \frac{\tilde{\pi}_{as}}{\tilde{\pi}_{Vs}}$:

$$\begin{aligned}
\ln(\tilde{\pi}_{as,1|V}) &= \rho_s \ln \tilde{v}_{as,1|V} - (\ln \tilde{\pi}_{Vs,1} - \ln \tilde{\pi}_{0s,1}) \\
\ln\left(\frac{\tilde{\pi}_{as,1}}{\tilde{\pi}_{Vs,1}}\right) &= \rho_s \ln \tilde{v}_{as,1|V} - (\ln \tilde{\pi}_{Vs,1} - \ln \tilde{\pi}_{0s,1}) \\
\ln(\tilde{\pi}_{as,1}) - \ln(\tilde{\pi}_{Vs,1}) + \ln \tilde{\pi}_{Vs,1} - \ln \tilde{\pi}_{0s,1} &= \rho_s \ln \tilde{v}_{as,1|V} \\
\frac{1}{\rho_s} \ln(\tilde{\pi}_{as,1}) - \ln \tilde{\pi}_{0s,1} &= \ln \tilde{v}_{as,1|V} \qquad (59)
\end{aligned}$$

Finally I substitute the expression for $\tilde{v}_{as|V}$ from Eq. (59) in Eq. (56) to get:

$$\frac{1-\delta}{\rho_s} \ln\left(\frac{\tilde{\pi}_{as,1}}{\tilde{\pi}_{0s,1}}\right) = q_{as} + \xi_{as} + \eta_{as} \qquad (60)$$

As, before I define the media forecast, $\tilde{\omega}^*_{as,1}$ as:

$$(1-\delta)\tilde{\omega}^*_{as,1} = q_{as} + \xi_{as} + \eta_{as} + \kappa_{sm} \qquad (61)$$

The next period updating, which can be iterated to t rounds of polling, can now be derived as:

$$\tilde{\vartheta}_{acs,2|\tilde{\omega}^*_{as,1}} = \tilde{\beta}_{a,2}[\tilde{\vartheta}_{acs,2} - \delta \ln(\tilde{v}_{acs,1|V})] + (1-\tilde{\beta}_{a,2})[(1-\delta)\tilde{\omega}^*_{as,1}] + \delta \ln(\tilde{v}_{acs,2|V}) \qquad (62)$$

I estimate the nesting parameter δ from the data based on Eq. (53) to determine the appropriate model structure. Note that the variable $\tilde{v}_{acs,1|v}$ is endogenous, suggesting the need for instruments that are correlated with the (within voters) alliance vote share. Berry (1994) recommends using the characteristics of other alternatives and number of alternatives as instruments. In line with this, I use the within state variation in the alliance composition, fraction of regional parties within an alliance, and the total number of parties contesting in a constituency as instruments. All the instruments are strongly correlated with the alliance (conditional) vote share, with a joint F-statistic of over 35. I calculate the unconditional seat share forecast by applying Bayes rule and multiplying the exit-poll forecast of turnout with the conditional seat share forecast ($\tilde{\pi}_{as} = \tilde{\pi}_{as|V} \times \tilde{\pi}_{Vs}$).

A.4 Data Sources

In this section I describe in detail the GIS data used for the analysis.

PC & AC Maps: The Parliamentary Constituency (PC), Assembly Constituency (AC) map, shape files were obtained from the Election Commission of India website (http://eci.gov.in/). This data was cleaned and geo-referenced using projections provided by Sandip Sukhtankar[58] and INRM Consultants, New Delhi. Further, the ECI provides images of the polling schedule for the 2004 elections which was converted into a raster and superimposed on the shape files. This enabled me to attach the exact polling phase and date to each PC and AC polygon. This procedure also allows me to digitize the exact polling phase boundary. Distances to the boundary point for each AC (the location identifier for each voter in the NES data) was

[58] Retrieved from http://www.dartmouth.edu/ sandip/data.html

calculated in two ways. First, I converted both the AC polygons and the boundary line into point locations – for the AC by taking the centroid of each polygon and for the boundary line by dividing the line into a set of points set apart by a distance of 7 km. I, then, calculate Euclidean distance between each pair of AC-Boundary point as well as take the nearest distance from each AC point to any point on the boundary line. Finally, I also divided the entire boundary line into segments (depending on their turn angle) and classify AC's according to the nearest boundary segment they belong to.

Rainfall: I use high resolution observed rainfall data provided by the India Meteorological Department (IMD) to calculate the amount of daily rainfall in each PC over the election period. Technical details of the rainfall data collection by IMD is given in Rajeevan, Bhate, Kale, and Lal (2006). The rainfall data is at a resolution of $0.5° \times 0.5°$ latitude by longitude grid points. Since a single point precipitation value is really not representative of the volume of precipitation falling over PC/AC area, precipitation values were converted to areal estimates using the Thiessen polygon method[59] (Thiessen 1911). The weather grids were superimposed on the PC/AC polygon for deriving the weighted means of the inputs for each of the PC/AC polygon. The centroid of each sub basin is then taken as the location for the weather station to be used for further analysis. Daily rainfall data is, therefore, aggregated as weighted sum.

Newspaper Circulation and Television Coverage: Town-wise newspaper circulation figures for all newspapers, weeklies, dailies and magazines in India were taken from the reports of the Audit Bureau of Circulation India[60]. These figures were then matched to the geo-spatial coordinates of each town in India as obtained from the Geonames website[61]. To map the newspaper circulation data to the AC/PCs, I first convert the AC/PC polygon to centroids and then calculate the minimum distance to the nearest town/city with the newspaper circulation data. Using these distances as weights, I aggregate all the circulation data for each PC/AC. I also supplement the newspapers data using the state-wise television coverage (percentage of total population) figures for the year 2004 provided by the national television agency, Doordarshan[62].

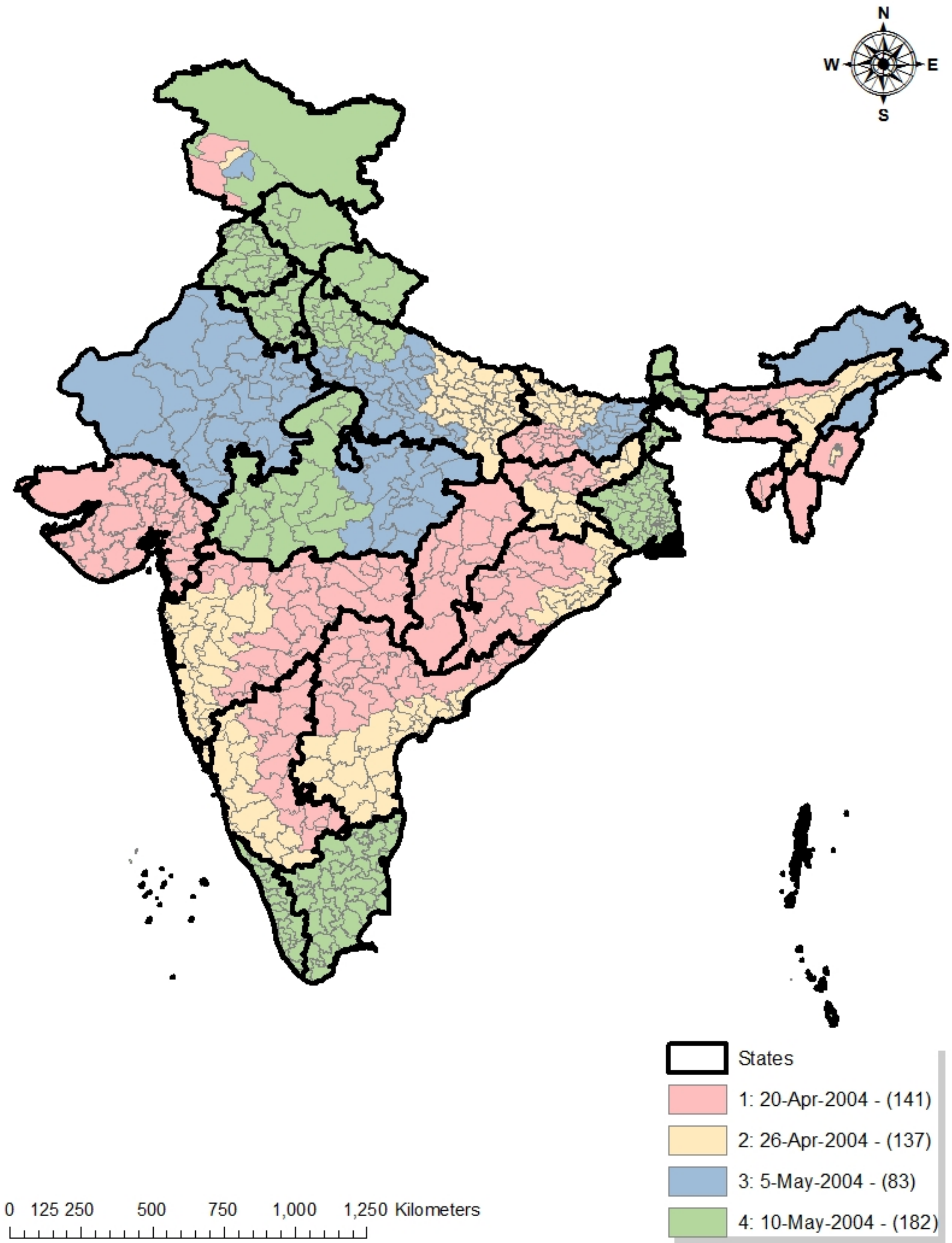
[59]The Thiessen polygon method is a graphical technique which calculates station weights based on the relative areas of each measurement station in the Thiessen polygon network. The individual weights are multiplied by the station observation and the values are summed to obtain the areal average precipitation.

[60]http://www.auditbureau.org/

[61]http://www.geonames.org. An approximate name matching algorithm was used to match town names in Geonames to town names in Audit Bureau report.

[62]Data is obtained through IndiaStat.

Figure 1: Election Schedule 2004 General Elections



Note: The map shows the polling sequence of all Parliamentary Constituencies across four pre-determined phases. The number of constituencies polling in every phase are indicated in parentheses, next to the polling date of each phase, in the legend.

Figure 2: Election Schedule 2004 General Elections: Exit poll Results After First Phase (20 April 2004)

Printed from
THE TIMES OF INDIA

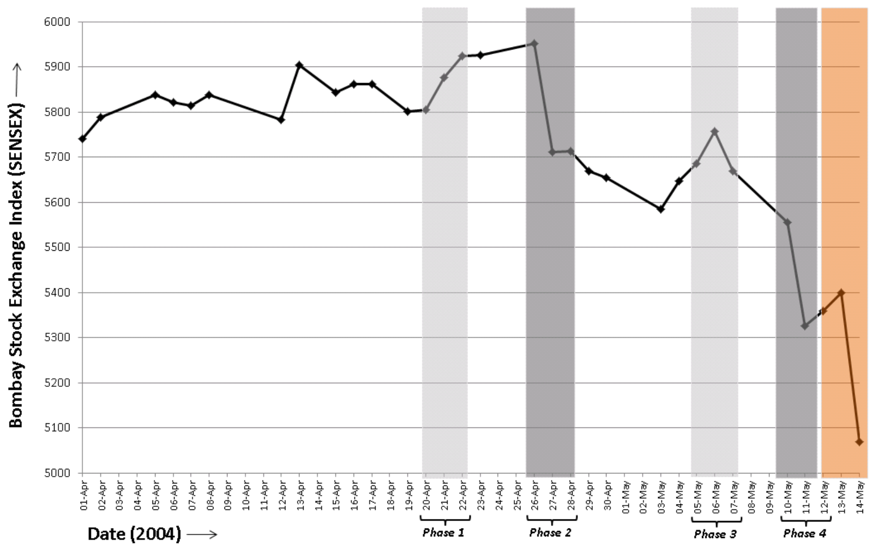
Exit Polls Phase 1

Apr 21, 2004, 11.30AM IST

	Sahara		Star News		Aaj Tak		Zee News *		NDTV		AVG **		TURNOUT (%)	
	NDA	CONG+	NDA	CONG+	NDA	CONG+	NDA	CONG+	NDA	CONG+	NDA	CONG+	2004	1999
AP (21)	11	9	7	13	6	15	10	11	3	17	7(-10)	13(+10)	55	69
MAHA (24)	10	13	7	17	11	12	11	7	15	8	11(-2)	12(+1)	43	67
GUJ (26)	21	5	22	4	23	3	19	7	18	8	21(+1)	5(-1)	48	47
CHHAT (11)	8	3	10	1	10	1	6	4	10	1	9(+1)	2(-1)	50	56
J'khand (6)	5	1	3	3	4	2	5	1	4	2	4(+0)	2(+0)	48	49
ASSAM (6)	1	3	2	3	4	1	1	2	1	2	2(+1)	2(-2)	60	73
K'taka (15)	9	5	10	4	11	4	4	5	11	3	9(+4)	4(-6)	58	65
BIHAR (11)	6	5	7	4	8	3	6	3	4	7	6(-2)	4(+1)	58	65
ORISSA (11)	7	4	9	2	9	2	7	4	8	3	8(-1)	3(+1)	53	57
OTHERS (9)	5	4	3	2	7	1	2	0	1	2	4(+0)	2(+2)	50	57
TOTAL (140)	83	52	80	53	93	44	71	43	75	53	80(-8)	49(+5)	53	60

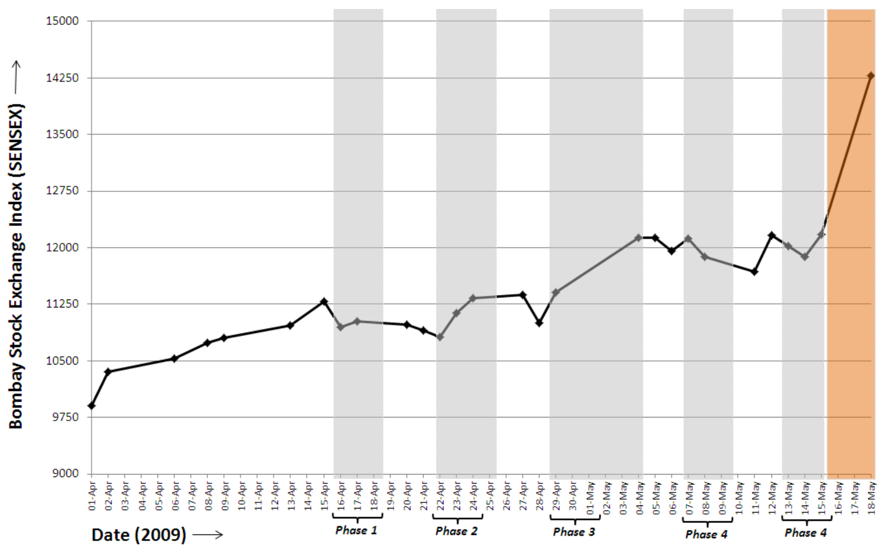
The figures in black brackets denote the number of constituencies polled on Tuesday.
* The figure for Zee News are averages of a projected range.
** Figures in red brackets indicate the change from the 1999 position.

Figure 3: BSE Index (SENSEX) over Election Period (April-May 2004)



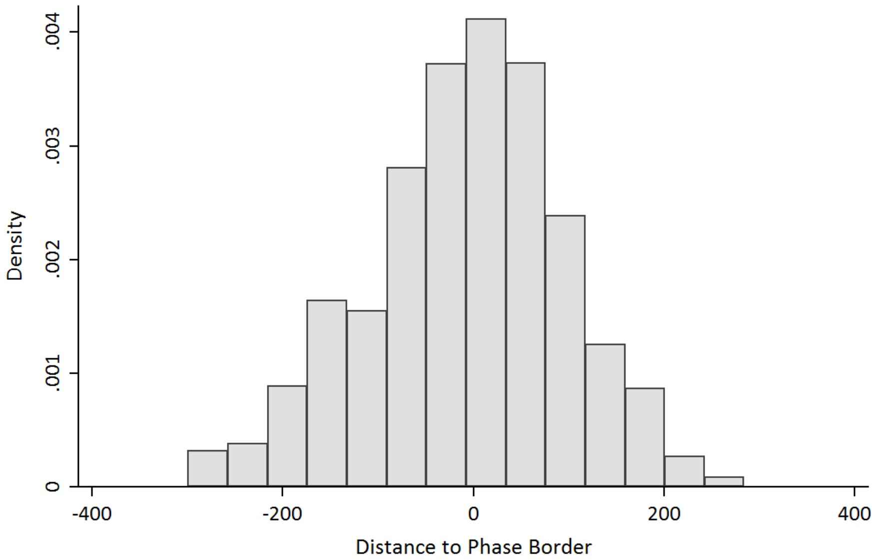
Note: The figure plots the daily trading index, SENSEX, over the election period, 1 April-14 May 2004. The blocks in grey highlights the +2/-2 daily window around each polling phase (with dark shades indicating larger fluctuations in the index). The block in orange highlights the +2/-2 daily window around the date when official results were finally declared.

Figure 4: BSE Index (SENSEX) over Election Period (April-May 2009)



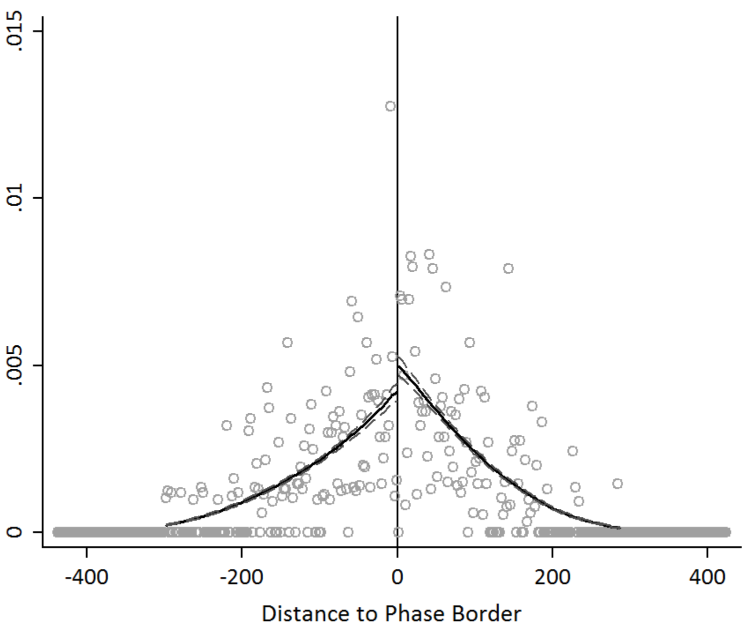
Note: The figure plots the daily trading index, SENSEX, over the election period, 1 April-18 May 2009. The blocks in grey highlights the +2/-2 daily window around each polling phase (with dark shades indicating larger fluctuations in the index). The block in orange highlights the +2/-2 daily window around the date when official results were finally declared.

Figure 5: Distance to Phase Boundary



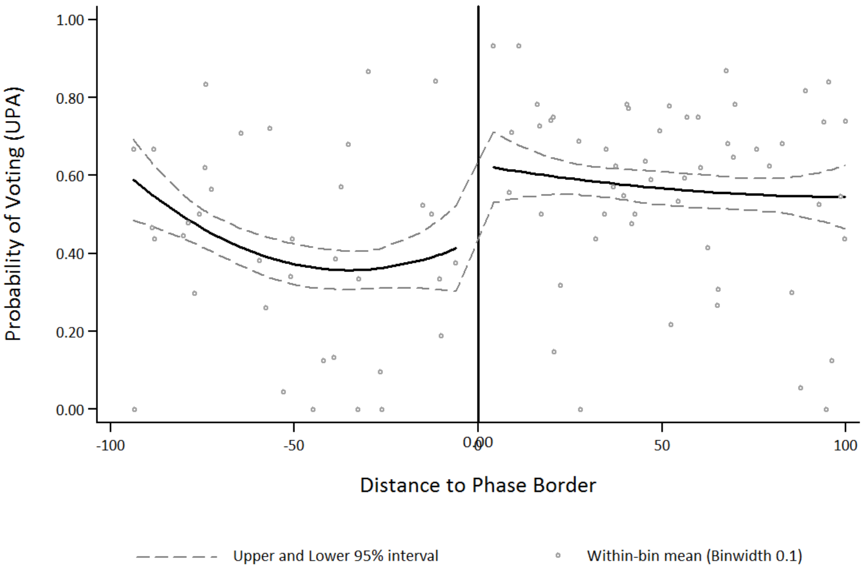
Note: The figure plots the histogram of each sample cluster's (AC) distance to the phase boundary. The distances are normalized, such that positive values indicate distances for those clusters that polled in Phase 2 (treatment) while negative values indicate distances for those clusters that polled in Phase 2 (control).

Figure 6: RD Validity: Density Smoothness Test for Distance to Phase Boundary



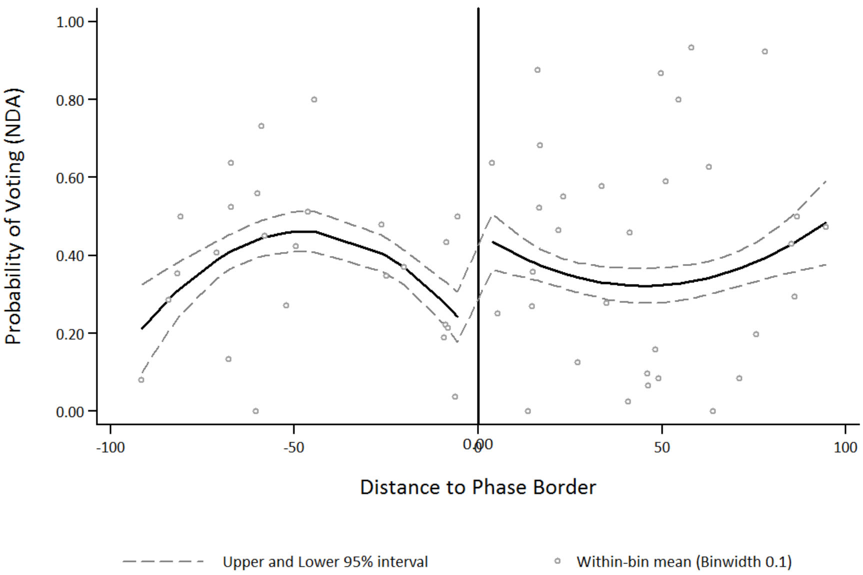
Note: The figure plots test for density smoothness proposed by McCrary (2008). The distances are normalized, such that positive values indicate distances for those clusters that polled in Phase 2 (treatment) while negative values indicate distances for those clusters that polled in Phase 2 (control).

Figure 7: Probability of Voting for UPA in UPA Gain States



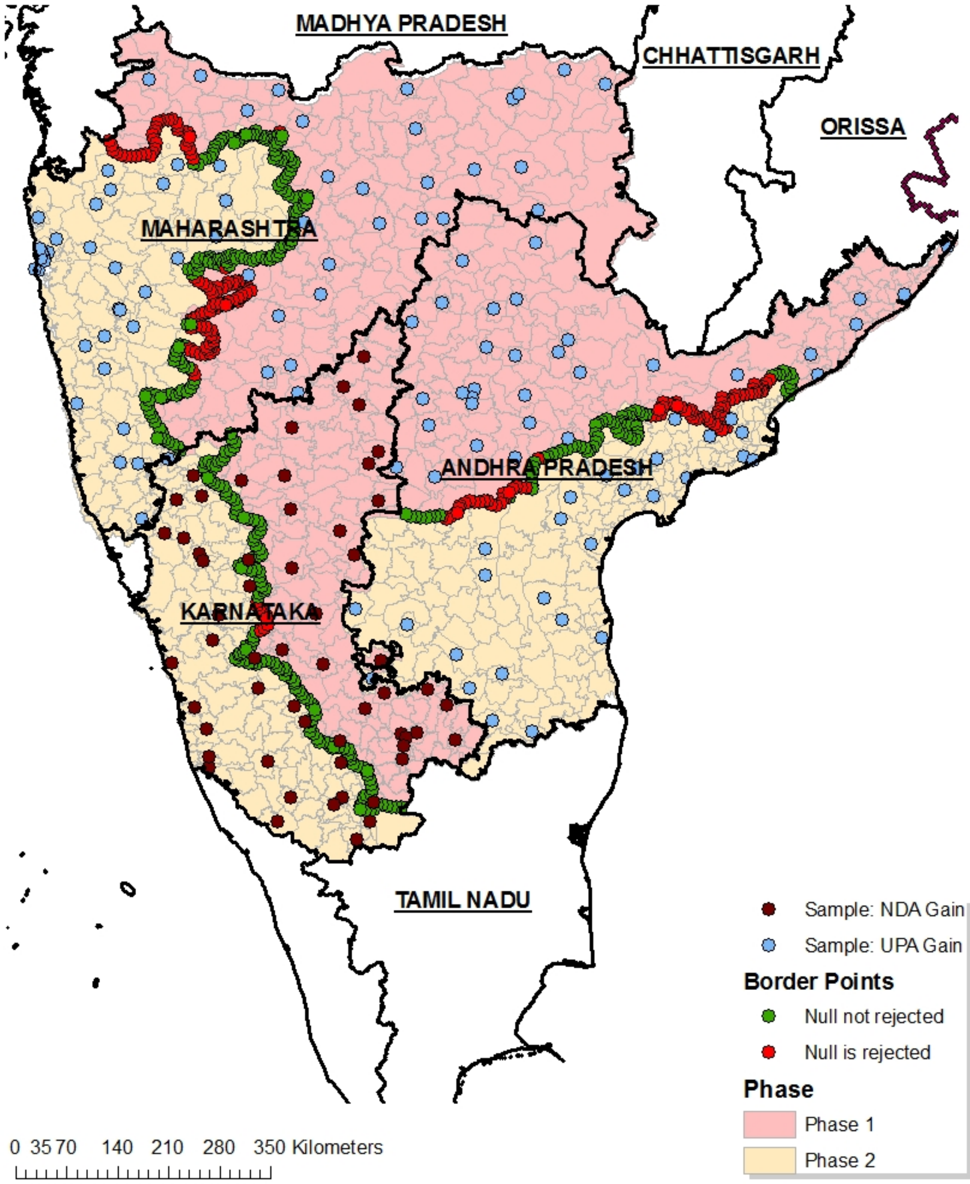
Note: The figure plots the local polynomial estimates of the probability of voting for the UPA in UPA gain states around the threshold distance. The distances are normalized, such that positive values indicate distances for those clusters that polled in Phase 2 (treatment) while negative values indicate distances for those clusters that polled in Phase 2 (control).

Figure 8: Probability of Voting for NDA in NDA Gain States



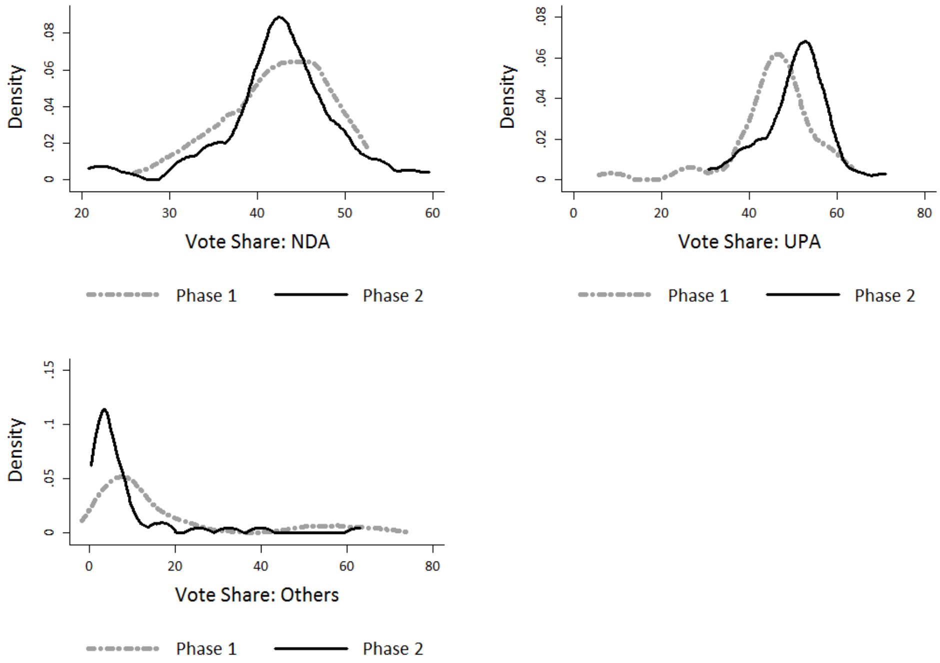
Note: The figure plots the local polynomial estimates of the probability of voting for the NDA in NDA gain states around the threshold distance. The distances are normalized, such that positive values indicate distances for those clusters that polled in Phase 2 (treatment) while negative values indicate distances for those clusters that polled in Phase 2 (control).

Figure 9: Geographical Discontinuity Estimates



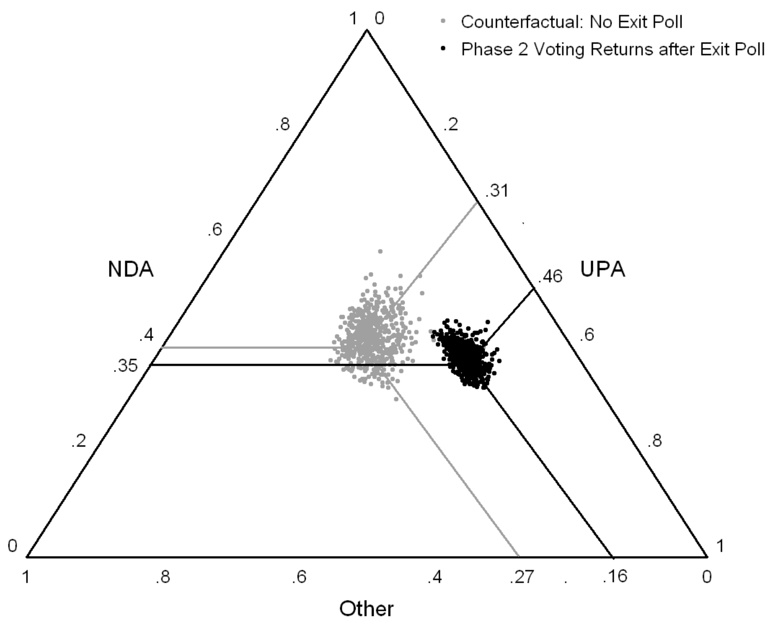
Note: The map plots the significance level of each treatment effect corresponding to points along the boundary line. All significant (at the 5% level) treatment effect estimates are coloured in red. Points located on either side of the boundary indicate sampled clusters.

Figure 10: Distribution of Actual Vote Shares



Note: The figure plots the kernel density distribution of each alliances's vote shares for Phase 1 (no exit polls) and Phase 2 constituencies.

Figure 11: Simulated Vote Shares



Note: The figure shows a ternary plot whose coordinates represent predicted fractions of the vote received by each alliance under different simulated election outcomes. Points coloured in black show the density distribution of simulated vote shares when voters receive exit poll results. Points coloured in gray show the density distribution of simulated vote shares when voters do not receive exit poll results.

Table 1: Summary Statistics: Aggregate, Constituency Level

	Phase 1		Phase 2		Phase 3		Phase 4		All	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
NDA:										
Vote Share	40.71	11.37	35.21	13.75	37.42	15.28	32.26	14.58	36.00	14.07
State Forecast	0.00	0.00	38.12	25.19	20.98	22.60	11.71	26.16	16.77	25.78
National Forecast	0.00	0.00	58.98	10.43	46.33	10.59	11.64	24.05	25.92	28.65
Vote Share 1999	39.41	13.90	38.39	15.85	35.65	14.99	31.94	19.04	36.09	16.66
Vote Share 1998	33.52	17.70	28.74	16.61	32.19	14.22	27.00	21.39	29.94	18.45
Assembly Vote Share	28.69	14.10	26.80	12.70	21.35	9.47	23.35	14.72	25.31	13.60
Others:										
Vote Share	19.73	19.75	26.39	25.85	35.93	27.69	34.55	27.08	28.83	25.90
State Forecast	0.00	0.00	4.86	8.04	17.69	22.60	5.03	15.14	5.62	14.17
National Forecast	0.00	0.00	4.58	1.03	9.97	3.70	2.65	7.53	3.57	5.60
Vote Share 1999	24.29	17.84	29.29	22.75	33.63	24.23	31.80	21.41	29.49	21.57
Vote Share 1998	20.64	20.40	28.45	22.90	35.21	22.15	33.51	21.41	29.13	22.30
Assembly Vote Share	43.82	19.72	50.80	22.62	63.77	15.52	49.79	20.23	50.63	20.99
UPA:										
Vote Share	39.57	13.12	38.40	17.40	26.66	17.45	33.19	20.87	35.17	18.21
State Forecast	0.00	0.00	29.97	21.22	11.55	14.88	2.64	5.50	10.24	17.40
National Forecast	0.00	0.00	32.42	7.80	38.66	11.03	4.91	10.31	15.78	18.01
Vote Share 1999	33.02	15.15	28.94	15.53	26.33	18.09	26.29	18.24	28.72	16.97
Vote Share 1998	32.63	18.17	26.48	21.07	23.74	21.87	22.46	17.38	26.33	19.66
Assembly Vote Share	24.27	11.67	21.60	13.67	13.81	9.82	20.54	14.93	20.75	13.48
Turnout	58.77	10.65	58.65	11.36	49.65	9.23	63.86	11.78	55.45	18.24
Turnout 1999	60.86	10.75	61.14	10.55	54.19	8.38	62.86	10.81	60.36	11.73
Turnout 1998	62.24	7.89	60.66	8.77	58.62	7.14	65.83	10.39	62.10	10.37
+1/-1 Rainfall (mm)	2.06	10.57	2.46	5.94	5.79	14.70	3.67	12.12		
+4/-4 Rainfall (mm)	1.53	3.82	1.99	2.45	2.74	7.65	2.35	6.43		

Table 2: Summary Statistics: NES Voter Post-Poll

	BW 100				BW150			
	Phase 1 (Control)		Phase 2 (Treatment)		Phase 1 (Control)		Phase 2 (Treatment)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Voted: NDA	0.42	0.49	0.37	0.48	0.41	0.49	0.39	0.49
Voted: UPA	0.39	0.49	0.48	0.50	0.43	0.50	0.46	0.50
Voted Others	0.19	0.39	0.15	0.36	0.16	0.36	0.16	0.36
Distance (in Kms)	47.17	26.67	46.47	27.16	66.29	41.29	63.78	42.55
Wealth	-0.31	2.05	-0.43	1.86	-0.27	2.04	-0.34	1.92
HH Size	6.74	3.27	6.54	3.24	6.66	3.24	6.50	3.18
Education	2.17	2.20	2.14	2.08	2.22	2.23	2.20	2.10
# Rooms	**3.00**	1.85	**2.82**	1.76	2.91	1.82	2.95	1.79
Hindu	0.81	0.39	0.82	0.39	0.81	0.39	0.81	0.39
Rural	0.22	0.41	0.19	0.39	0.23	0.42	0.20	0.40
Votercard	0.55	0.50	0.58	0.49	0.58	0.49	0.58	0.49

Table 3: Estimated Updating Weights (Phase 1 & 2): State Signal

	NDA			UPA		
	Coefficient	95% C.I.		Coefficient	95% C.I.	
Prior	0.801***	0.581	1.019	0.631***	0.265	0.996
State Signal	0.199*	-0.019	0.418	0.368**	0.003	0.734
$\Delta_{c,s}^{a,\downarrow}$	1.785	1.365	2.205	1.964	1.406	2.522
$\Delta_{c,s}^{a,\uparrow}$	0.755§	0.622	0.888	0.531§	0.321	0.742
N	96			96		
R2	0.763			0.769		

Notes: 1. $Prior$ is the predicted log-odds of vote share for Phase 2 constituencies. Predicted values are generated using the estimated parameters from a first stage regression of log-odds of vote share on state fixed effects, log-odds of vote share in the previous two elections and log-odds of vote share in the assembly elections over the sample of Phase 1 constituencies. The R2 for the respective first stages are 0.63 (NDA) and 0.54 (UPA); first stage parameter estimates are not reported but are available upon request. $StateSignal$ is the log-odds of state seat share as predicted by the exit-polls and normalized by the state swing factors (reported in the appendix). $\Delta_{c,s}^{a,\uparrow}$ indicates positive surprise while $\Delta_{c,s}^{a,\downarrow}$ indicates negative surprise. 2. Standard errors (in parentheses) are nonparametrically bootstrapped and normal confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 4: Estimated Updating Weights (Phase 1 & 2) : State and National Signals

	NDA			UPA		
	Coefficient	95% C.I.		Coefficient	95% C.I.	
Prior	0.801***	0.559	1.041	0.631***	0.270	0.991
State Signal	0.199*	-0.022	0.421	0.368**	0.009	0.728
National Signal	0	-0.120	0.120	0	-0.047	0.047
N	96			96		
R2	0.810			0.829		
$\sigma_{a,t}^2$ (Prior variance)	0.1509			0.170		

Notes: 1. $Prior$ is the predicted log-odds of vote share for Phase 2 constituencies. Predicted values are generated using the estimated parameters from a first stage regression of log-odds of vote share on state fixed effects, log-odds of vote share in the previous two elections and log-odds of vote share in the assembly elections over the sample of Phase 1 constituencies. The R2 for the respective first stages are 0.63 (NDA) and 0.54 (UPA); first stage parameter estimates are not reported but are available upon request. $StateSignal$ is the log-odds of state seat share as predicted by the exit-polls and normalized by the state swing factors (reported in the appendix). $NationalSignal$ is the log-odds of average seat share across all states polling in Phase 1 as predicted by the exit-polls and normalized by the average national swing factor. $\sigma_{a,t}^2$ is the within-state variance in PC fixed effects, estimated from individual post-poll data. 2. Standard errors (in parentheses) are nonparametrically bootstrapped and normal confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 5: Estimated Updating Weights (Phase 3 & 4): State and National Signals

	NDA			UPA		
	Coefficient	95% C.I.		Coefficient	95% C.I.	
Prior	0.647^{***}	0.178	1.116	0.779^{***}	0.435	1.123
State Signal	0.352	-0.165	0.870	0.368	-0.067	0.508
National Signal	0	-0.205	0.205	0	-0.378	0.378
N	30			30		
R2	0.753			0.808		

Notes: 1. $Prior$ is the predicted log-odds of vote share for Phase 4 constituencies. Predicted values are generated using the estimated parameters from a first stage regression of log-odds of vote share on state fixed effects, log-odds of vote share in the previous two elections and log-odds of vote share in the assembly elections over the sample of Phase 3 constituencies. The R2 for the respective first stages are 0.88 (NDA) and 0.84 (UPA); first stage parameter estimates are not reported but are available upon request. $StateSignal$ is the log-odds of state seat share as predicted by the exit-polls and normalized by the state swing factors (reported in the appendix). $NationalSignal$ is the log-odds of average seat share across all states polling in Phase 1 as predicted by the exit-polls and normalized by the average national swing factor. 2. Standard errors (in parentheses) are nonparametrically bootstrapped and normal confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 6: Assembly Elections (Phase 1 & 2): State and National Signals

	NDA			UPA		
	Coefficient	95% C.I.		Coefficient	95% C.I.	
Prior	0.946^{***}	0.868	1.024	0.951^{***}	0.679	1.22
State Signal	0.053	-0.017	0.123	0.048	-0.064	0.161
National Signal	0	0	0.000	0	0	0.201
N	443			443		
R2	0.921			0.707		

Notes: 1. $Prior$ is the predicted log-odds of assembly (state election) vote share for Phase 2 constituencies. Predicted values are generated using the estimated parameters from a first stage regression of log-odds of assembly vote share on state fixed effects, log-odds of vote share in the previous two general elections and log-odds of vote share in the assembly elections over the sample of Phase 1 constituencies. The R2 for the respective first stages are 0.90 (NDA) and 0.68 (UPA); first stage parameter estimates are not reported but are available upon request. $StateSignal$ is the log-odds of state seat share (general elections) as predicted by the exit-polls and normalized by the state swing factors (reported in the appendix). $NationalSignal$ is the log-odds of average seat share (general elections) across all states polling in Phase 1 as predicted by the exit-polls and normalized by the average national swing factor. 2. Standard errors (in parentheses) are nonparametrically bootstrapped and normal confidence intervals (based on bootstrap estimates) are reported. They are corrected for clustering within a PC using the Moulton factor (1.9). 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 7: Voter Turnout (Phase 1 & 2)

	NDA			UPA		
	Coefficient	90% C.I.		Coefficient	90% C.I.	
Prior	0.792***	0.458	1.241	0.762***	0.368	0.990
State Signal	0.207	-0.241	0.541	0.237*	0.009	0.631
N	96			96		
R2	0.614			0.555		

Notes: 1. $Prior$ is the predicted value of $\widetilde{\vartheta}_{acs,2} - \delta \ln(\widetilde{\upsilon}_{acs,1|V})$ as derived in Eq. (62). Predicted values are generated using the estimated parameters from a first stage regression of log-odds of $\widetilde{\vartheta}_{acs,2} - \delta \ln(\widetilde{\upsilon}_{acs,1|V})$ on state fixed effects and log-odds of vote share in the previous two general elections over the sample of Phase 1 constituencies. δ is estimated from a regression of overall log-odds of (unconditional) vote shares on control variables and number of competing parties within states and within an alliance as instruments (with a bootstrapped F-stat of 36.4). As before the two first stage parameter estimates are not reported but are available upon request. $StateSignal$ is the log-odds of state seat share (general elections) times the turnout as predicted by the exit-polls and normalized by the state swing factors (reported in the appendix). 2. Standard errors (in parentheses) are nonparametrically bootstrapped and percentile confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 8: Selection Correction (Phase 1 & 2)

Selection Stage Estimates			
-4/+4 Rainfall	-0.140***	-0.17	-0.10
-1/0 Rainfall	0.093***	0.035	0.115

	NDA			UPA		
	Coefficient	90% C.I.		Coefficient	90% C.I.	
Prior	0.771***	0.349	0.919	0.666***	0.188	0.942
State Signal	0.228*	0.080	0.650	0.333*	0.057	0.811
Mills Ratio	-0.278***	-0.415	-0.102	-0.384**	-0.515	-0.238
N	96			96		
R2	0.805			0.836		

Notes: 1. $Prior$ and $StateSignal$ are defined as before in Table 3. $MillsRatio$ is the selection correction term from a first stage conditional logit regression of the probability of polling in a particular phase on the +4/-4 and +1/-1 log of daily rainfall window around polling date. 2. Standard errors (in parentheses) are nonparametrically bootstrapped and percentile confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 9: Selection Correction & Voter Turnout (Phase 1 & 2)

	NDA			UPA		
	Coefficient	90% C.I.		Coefficient	90% C.I.	
Prior	0.8069^{***}	0.411	1.201	0.8067^{***}	0.382	0.993
State Signal	0.193	-0.201	0.588	0.193^{*}	0.006	0.617
Mills Ratio	-0.077	-0.174	0.003	-0.067^{*}	-0.211	-0.011
N	96			96		
R2	0.618			0.617		

Notes: 1. $Prior$ and $StateSignal$ are defined as before in Table 7. $MillsRatio$ is the selection correction term from a first stage conditional logit regression of the probability of polling in a particular phase on the +4/-4 and +1/-1 log of daily rainfall window around polling date. 2. Standard errors (in parentheses) are nonparametrically bootstrapped and percentile confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 10: Media Weighted State Signals

	NDA			UPA		
	Coefficient	95% C.I.		Coefficient	95% C.I.	
Prior	0.779^{***}	0.603	0.954	0.590^{***}	0.291	0.889
State Signal	0.2209^{**}	0.045	0.396	0.409^{**}	0.110	0.708
N	443			443		
R2	0.701			0.704		

Notes: 1. $Prior$ and $StateSignal$ are defined as before in Table 3. 2. The regression is weighted using a kernel function of newspaper and television coverage in each PC. 4. Standard errors (in parentheses) are nonparametrically bootstrapped and normal confidence intervals (based on bootstrap estimates) are reported. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 11: Regression Discontinuity Estimates: Voting for UPA in UPA Gain States

	Latitude-Longitude					
	BW 100	BW 150	BW 200	BW 100	BW 150	BW 200
Phase 2	0.198**	0.235***	0.208**	0.189*	0.188*	0.190*
	(0.089)	(0.085)	(0.085)	(0.105)	(0.120)	(0.104)
Pseudo-R2	0.120	0.114	0.101	0.196	0.159	0.146
Lat/Long Polynomials	Yes	Yes	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes	Yes	Yes
Lat/Long × Boundary-Segment F.E	No	No	No	Yes	Yes	Yes
	Distance					
	BW 100	BW 150	BW 200	BW 100	BW 150	BW 200
Phase 2	0.255**	0.246***	0.248***	0.322***	0.295***	0.350***
	(0.114)	(0.093)	(0.084)	(0.103)	(0.081)	(0.070)
Pseudo-R2	0.119	0.106	0.100	0.116	0.105	0.100
Distance Polynomials	Yes	Yes	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes	Yes	Yes
Distance × Boundary-Segment F.E	No	No	No	Yes	Yes	Yes
N	1757	2295	2654	1757	2295	2654
# Clusters (AC)	94	124	145	94	124	145

Notes: 1. *Phase*2 is the binary treatment variable, taking the value one for all voters who polled in Phase 2 of the elections. *BW* is Bandwidth. Latitude (x) and Longitude (y) polynomials are: $x, y, xy, x^2, y^2, x^3, y^3, x^2y, y^2x$; Distance polynomials are cubics of distance to the boundary. Other control variables include: education, income, wealth, household size, number of rooms, religion, rural/urban indicator. 2. Standard errors (in parentheses) are adjusted for clustering at the Assembly Constituency (AC) level. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 12: Regression Discontinuity Estimates: Voting for NDA in NDA Gain States

	Latitude-Longitude					
	BW 100	BW 150	BW 200	BW 100	BW 150	BW 200
Phase 2	-0.057	-0.112	-0.073	-0.047	-0.135*	-1.384*
	(0.491)	(0.219)	(0.283)	(0.077)	(0.074)	(0.080)
Pseudo-R2	0.185	0.148	0.148	0.197	0.148	0.156
Lat/Long Polynomials	Yes	Yes	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes	Yes	Yes
Lat/Long × Boundary-Segment F.E	No	No	No	Yes	Yes	Yes
	Distance					
	BW 100	BW 150	BW 200	BW 100	BW 150	BW 200
Phase 2	-0.000	-0.067	-0.088	-0.107	-0.179**	-0.209***
	(0.058)	(0.066)	(0.062)	(0.079)	(0.080)	(0.075)
Pseudo-R2	0.157	0.128	0.138	0.134	0.105	0.118
Distance Polynomials	Yes	Yes	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes	Yes	Yes
Distance × Boundary-Segment F.E	No	No	No	Yes	Yes	Yes
N	1802	2294	2661	1802	2294	2661
# Clusters (AC)	59	74	84	59	74	84

Notes: 1. *Phase*2 is the binary treatment variable, taking the value one for all voters who polled in Phase 2 of the elections. *BW* is Bandwidth. Latitude (x) and Longitude (y) polynomials are: $x, y, xy, x^2, y^2, x^3, y^3, x^2y, y^2x$; Distance polynomials are cubics of distance to the boundary. Other control variables include: education, income, wealth, household size, number of rooms, religion, rural/urban indicator. 2. Standard errors (in parentheses) are adjusted for clustering at the Assembly Constituency (AC) level. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 13: Regression Discontinuity Estimates: Narrow Bandwidths

	Latitude-Longitude			
	UPA		NDA	
	BW 50	BW 75	BW 50	BW 75
Phase 2	0.414***	0.214**	-0.041	0.024
	(0.117)	(0.098)	(0.069)	(0.39)
Pseudo-R2	0.228	0.148	0.188	0.203
Lat/Long Polynomials	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes
	Distance			
	UPA		NDA	
	BW 50	BW 75	BW 50	BW 75
Phase 2	0.426**	0.244*	0.106	0.031
	(0.172)	(0.131)	(0.072)	(0.068)
Pseudo-R2	0.214	0.144	0.220	0.201
Distance Polynomials	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes
N	809	1348	1090	1519
# Clusters (AC)	45	71	34	49

Notes: 1. *Phase*2 is the binary treatment variable, taking the value one for all voters who polled in Phase 2 of the elections. *BW* is Bandwidth. Latitude (x) and Longitude (y) polynomials are: $x, y, xy, x^2, y^2, x^3, y^3, x^2y, y^2x$; Distance polynomials are cubics of distance to the boundary. Other control variables include: education, income, wealth, household size, number of rooms, religion, rural/urban indicator. 2. Standard errors (in parentheses) are adjusted for clustering at the Assembly Constituency (AC) level. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 14: Regression Discontinuity Estimates: Placebo (Voting for UPA)

	Latitude-Longitude					
Voting for:	Assembly Elections			1999 General Elections		
	BW 100	BW 150	BW 200	BW 100	BW 150	BW 200
Phase 2	-0.094	-0.081	-0.105	-0.564**	-0.295	-0.190
	(0.123)	(0.098)	(0.096)	(0.218)	(0.293)	(0.182)
Pseudo-R2	0.103	0.081	0.064	0.087	0.071	0.062
Lat/Long Polynomials	Yes	Yes	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes	Yes	Yes
	Distance					
Voting for:	Assembly Elections			1999 General Elections		
	BW 100	BW 150	BW 200	BW 100	BW 150	BW 200
Phase 2	-0.011	-0.028	0.039	-0.013	0.070	0.171
	(0.119)	(0.096)	(0.088)	(0.167)	(0.184)	(0.158)
Pseudo-R2	0.106	0.066	0.054	0.079	0.063	0.060
Distance Polynomials	Yes	Yes	Yes	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes	Yes	Yes	Yes
N	1352	1788	2137	1784	2322	2681
# Clusters (AC)	62	84	104	94	124	145

Notes: 1. $Phase2$ is the binary treatment variable, taking the value one for all voters who polled in Phase 2 of the elections. BW is Bandwidth. Latitude (x) and Longitude (y) polynomials are: $x, y, xy, x^2, y^2, x^3, y^3, x^2y, y^2x$; Distance polynomials are cubics of distance to the boundary. Other control variables include: education, income, wealth, household size, number of rooms, religion, rural/urban indicator. 2. Standard errors (in parentheses) are adjusted for clustering at the Assembly Constituency (AC) level. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 15: Regression Discontinuity Estimates: Voter Turnout Estimates

	Voting for UPA in UPA Gain States		
	BW 100	BW 150	BW 200
Phase 2 (Log-Odds)	0.960***	0.849***	0.773***
	(0.321)	(0.250)	(0.233)
Predicted Change in Voting Choice (Conditional)‡	0.180	0.168	0.159
Predicted Change in Voter Turnout†	0.319	0.276	0.246
δ (Nesting Parameter)	0.40*	0.40*	0.40*
	(0.311)	(0.311)	(0.311)
Distance Polynomials	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes
N	1757	2295	2654
# Clusters (AC)	94	124	145

	Voting for NDA in NDA Gain States		
	BW 100	BW 150	BW 200
Phase 2 (Log-Odds)	0.059	-0.174	-0.265
	(0.157)	(0.189)	(0.197)
Predicted Change in Voting Choice (Conditional)‡	0.092	0.078	0.073
Predicted Change in Voter Turnout†	-0.014	-0.079	-0.101
δ (Nesting Parameter)	0.40*	0.40*	0.40*
	(0.311)	(0.311)	(0.311)
Distance Polynomials	Yes	Yes	Yes
Boundary-Segment F.E	Yes	Yes	Yes
N	1802	2294	2661
# Clusters (AC)	59	74	84

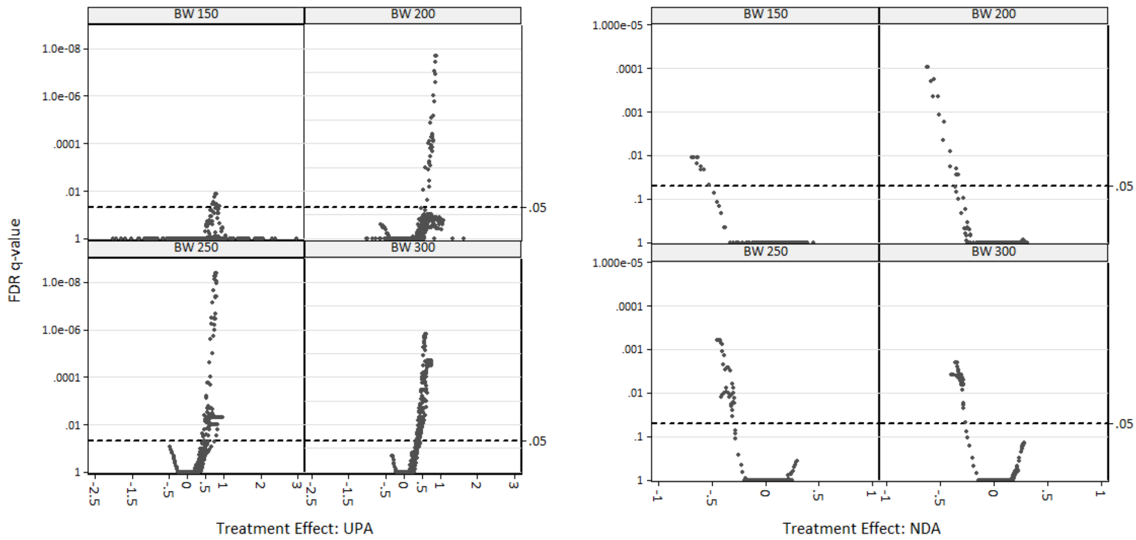
Notes: 1. *Phase*2 is the binary treatment variable, taking the value one for all voters who polled in Phase 2 of the elections. *BW* is Bandwidth. Distance polynomials are squares of distance to the boundary. Other control variables include: education, income, wealth, household size, number of rooms, religion, rural/urban indicator. 3. ‡ is the change in probability of choosing an alliance conditional on the decision to vote; † is the change in probability of voting. Both variables measure the predicted change in probability when the treatment variable (*Phase*2) switches from zero to one, holding all control variables at their mean values. 3. Standard errors (in parentheses) are adjusted for clustering at the Assembly Constituency (AC) level. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Table 16: Exit Polls and Voter Turnout

	(1) Pr(Voting)	(2) Pr(Voting)	(3) Pr(Voting)	(4) Pr(Voting)
UPA Surprise	0.718*			
	(0.369)			
NDA Surprise	0.490			
	(0.348)			
UPA Lead		1.181		
		(0.972)		
NDA Lead		0.388		
		(0.365)		
UPA Narrow Win			0.771**	
			(0.383)	
NDA Narrow Win			-0.055	
			(0.394)	
UPA Seat Share Forecast				1.658
				(1.085)
NDA Seat Share Forecast				-0.500
				(1.002)
UPA Voter (1999)	1.402**	1.636**	1.743**	1.379**
	(0.187)	(0.162)	(0.192)	(0.180)
× UPA Surprise	-0.122			
	(0.287)			
× NDA Surprise	0.094			
	(0.375)			
× UPA Lead		0.270		
		(0.514)		
× NDA Lead		0.116		
		(0.367)		
× UPA Narrow Win			0.360	
			(0.379)	
× NDA Narrow Win			-0.576	
			(0.381)	
× UPA Seat Share Forecast				-0.003
				(1.237)
× NDA Seat Share Forecast				0.040
				(1.207)
NDA Voter (1999)	1.322**	1.519**	1.624**	1.297**
	(0.175)	(0.159)	(0.180)	(0.172)
× UPA Surprise	-0.160			
	(0.261)			
× NDA Surprise	0.166			
	(0.354)			
× UPA Lead		0.337		
		(0.562)		
× NDA Lead		0.020		
		(0.290)		
× UPA Narrow Win			0.145	
			(0.372)	
× NDA Narrow Win			-0.378	
			(0.373)	
× UPA Seat Share Forecast				0.031
				(1.101)
× UPA Seat Share Forecast				-0.068
				(0.976)
N	5180	4469	4469	5180

Notes: 1. *Surprise* is the interaction of forecasted 'gain' and *Phase*2 (treatment variable); *Lead* is the interaction of forecasted seat share greater than half and *Phase*2 (treatment variable); *Narrow Win* is the interaction of forecasted seat share greater than one-third and *Phase*2 (treatment variable);*Seat Share (Forecast)* is the interaction of forecasted seat share and *Phase*2 (treatment variable).
All resuls are estimates at the 150 Km Bandwidth with 152 AC clusters and include distance polynomials (squares of distance to the boundary), education, income, wealth, household size, number of rooms, religion, rural/urban indicator. 3. Standard errors (in parentheses) are adjusted for clustering at the Assembly Constituency (AC) level. 3. $*$ indicates significance at 10%; ** at 5%; *** at 1%.

Figure 12: Geographical Discontinuity Estimates: Treatment Effects



Note: The figure plots the the distribution of coefficient estimates (treatment effects), corresponding to each boundary point, against its q-values (corrected critical p-values). The figure shows the subset of p-values that are at or below the corrected critical p-value under any FDR. There is one panel for each bandwidth used. The horizontal axis gives the coefficient on the treatment effect and the vertical axis gives the corresponding q-value on a reverse log scale. The horizontal reference line (in dashed black line) indicates the 95% confidence level. All points above this line are statistically significant at 5%.

Table 17: Geographical Discontinuity Estimates: Average Treatment Effect

	Obs.	Average # of AC Clusters	Treatment Effect at: 90% Confidence	95% Confidence	99% Confidence
UPA Gain States					
BW 150	411	12.66	0.06	0.04	0.00
BW 200	411	20.66	0.08	0.07	0.07
BW 250	411	29.85	0.20	0.18	0.15
BW 300	411	39.39	0.26	0.23	0.20
BW 350	411	46.51	0.26	0.25	0.22
NDA Gain States					
BW 150	150	17.49	-0.04	-0.04	0.00
BW 200	150	25.60	-0.06	-0.06	-0.05
BW 250	150	33.36	-0.08	-0.08	-0.05
BW 300	150	39.93	-0.08	-0.08	-0.07
BW 350	150	45.22	-0.07	-0.08	-0.07

Notes: 1. This table reports the mean value of treatment effects estimated for each point on the boundary and for each bandwidth. All treatment effects that have a q-value greater than 0.05 are set to zero. *BW* is Bandwidth. The treatment effect at each boundary point is obtained by regressing the probability of voting for UPA/NDA on distance polynomials and the treatment (Phase 2) variable.