THE DATA STORY

Uttar Pradesh
Elections 2022
— Phase 1

CPR CENTRE FOR
POLICY
RESEARCH

The Data Story: Uttar Pradesh Elections 2022 – Phase 1

DataCommons @ CPR

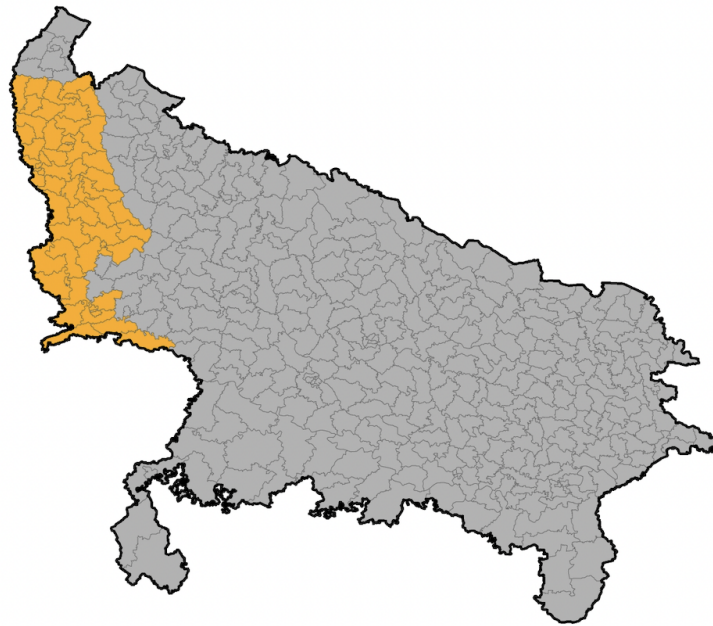10 February 2022

1 Introduction

Figure 1: Phase 1 Assembly Constituencies

In the first phase of the Uttar Pradesh elections of 2022, 58 assembly constituencies (ACs) will go to the polls. Polling will take place on 10 February, 2022. This short piece constitutes the first piece in our data collaboration with *Datalok*,[1] who have provided detailed, digitized data for every polling booth in Uttar Pradesh (UP) over multiple elections. Our analysis here makes use of electoral results in the 2017 election disaggregated into the 21,271 polling booths nested within the 58 ACs going to the polls.

Three key lessons emerge from this exercise.

---

[1]Datalok is a research and analysis firm that helps organizations take data-driven informed decisions. For more details, please visit datalok.in.

First, the Muzaffarnagar riots of 2013, and subsequent Hindu-Muslim polarization, transformed the competitive multiparty nature of the party system into a stable area of support for the Bharatiya Janata Party (BJP). In particular, religious polarization created a narrative frame to generate caste consolidations and "countermobilization" against certain politically significant communities, namely Muslims and Jatavs, in the region.

Second, this pattern of Hindu-Muslim polarization structured local politics (polling booth level) in key ways. With a successful countermobilzation strategy, the BJP built stable, large majorities of local support in which the election was largely uncompetitive. Its competitors, the Samajwadi Party (SP), the Bahujan Samaj Party (BSP), Rashtriya Lok Dal (RLD) and the Congress Party (INC), were limited to support in concentrated clusters of the Muslim, Jat and Jatav community alongside a small base of loyal party supporters.

Third, the areas in which BJP's countermobilization strategy was less successful saw fractionalization among voters to yield multiparty or fragmented party competition at the local level. Such fragmentation effectively reduced the vote share threshold a party must reach to win the polling booth. In such a scenario, locally numerous and dominant communities, e.g., Muslims, Jats, Jatavs, could consolidate to make caste-specific parties victorious.

In this note, we describe the historical empirical pattern of voting in these 58 ACs, with some concluding thoughts on how recent events and ground-level observations may affect these dynamics. Section 2 describes recent events in the region alongside demographic factors to understand the outcomes of the 2012 and 2017 state elections, and the 2014 and 2019 national elections, across these 58 ACs. Section 3 delves deeper into the polling booth trends from the 2017 election. Section 4 provides brief concluding thoughts, with some insights into what we see on the ground today.

2   Background and Context

The traditional demographic context of the Phase 1 areas is that of multiple caste assertions and a resulting party fragmentation due to caste-based parties. The 58 ACs in the first phase correspond to areas contained in the districts of Aligarh, Agra, Baghpat, Bulandshahr, Gautam Buddha Nagar, Ghaziabad, Mathura, Meerut, Muzaffarnagar, Panchsheel Nagar and Prabudh Nagar (the latter two districts were created after the 2011 Census, carved out from Ghaziabad and Muzaffarnagar, respectively). To provide some context, we contextualize the bases of various parties with Indian Census data. The percentage of Muslims, scheduled castes (SCs), and the percentage of Jatavs among SCs is shown for these districts in figure 2.

Mayawati, the leader of the BSP, is from this region (Gautam Buddha Nagar). The BSP is known to appeal particularly to the Jatav community – the largest community among SCs in UP at 54%. The areas in phase 1 correspond to districts with a particularly high percentage of Jatavs. According to the 2011 Indian Census, the percentage Jatavs of among SCs is at least 70% in almost every district in this phase, much higher than the UP-wide percentage. It is interesting to note that the percentage of SCs in this region is not unusually high for UP (with an SC population of 23% across UP according to the 2011 Census). Rather, it is the high concentration of Jatavs

among SCs that has traditionally made this region a BSP stronghold.

The Samajwadi Party on the other hand, which is viewed as particularly catering to the Yadav community and the Muslim community, has a weaker base in these areas. While we do not have official caste data, the Yadav community is visibly less numerous in phase 1 areas as compared to much of UP. There is a noticeably large Muslim population in many districts in this region. According to the 2011 Indian Census, 6 out of 9 phase 1 districts have larger Muslim populations than the UP-wide average of 19%. While there are significant Muslim populations in much of the region, in practice a significant portion of the Muslim community in the area has strategically supported the BSP given its numerical strength in the past.

This region also has a vibrant agricultural economy, with noticeable specialization in sugarcane. The Jat community is numerous in this region (although once again we do not have official data), and is a community that has traditionally been linked with the welfare of agriculture here. This Jat-farmer alliance has been crucial in building support for the RLD, a party started by Chaudhary Ajit Singh, son of former Prime Minister Chaudhary Charan Singh. It is today headed by his son, Jayant Chaudhary. It also seen as one of the earliest parties to strike at a stable base of Congress support across India. However, in recent decades the RLD had seen an erosion of support, particularly outside the Jat community.

Districts	Muslim Population (%)	SC Population (%)	Jatav Population of SCs (%)
Agra	9	22	70
Aligarh	20	21	59
Baghpat	28	11	68
Bulandshahr	22	21	75
Gautam Buddha Nagar	13	6	65
Ghaziabad	25	17	71
Mathura	9	20	72
Meerut	34	18	75
Muzaffarnagar	41	14	75
Total	23	18	

Figure 2: Demographics of Phase 1 Districts

These data show percentages of the Muslim and SC communities in the districts corresponding to the 58 ACs going to the polls in phase 1, according to the 2011 Indian Census. Note that two districts, Panchsheel Nagar and Prabudh Nagar, were carved out from Ghaziabad and Muzaffarnagar, respectively after the Census. As such, the data corresponding to these new districts is included in the census figures for the parent districts.

In September 2013, Muzaffarnagar district saw brutal Hindu-Muslim riots that, according to government data, claimed the lives of more than 60 people and displaced tens of thousands more. The riots purportedly arose from local conflict between the Jat community and the Muslim community, and the resulting Hindu-Muslim polarization transformed the politics of the region

and bases of party support. In particular, this event precipitated a wholesale shift of many Hindu voters to the BJP – even among communities, like the Jat community, that had not historically supported the BJP.

Party	2012	2014	2017	2019
BJP	10	57	53	46
BSP	20	0	2	7
INC	5	0	0	0
RLD	9	0	1	4
SP	14	1	2	1

Figure 3: Seat Share in Phase 1 ACs in Recent Elections

This table displays the seat shares of parties for the 4 most recent elections across the 58 ACs in phase 1. The data for 2012 and 2017 correspond to the winning party for the Vidhan Sabha seat in state elections, whereas the data for 2014 and 2019 correspond to the leading party in the AC in the national election. These data are derived from publicly released documents from the Election Commission of India (ECI).

Figure 3 displays the seat shares in the 2012 and 2017 state elections, and the leading parties in the 2014 and 2019 national elections, across the 58 ACs in phase 1. Until 2012, this region displayed a fragmented party system, in which caste-specific and community-specific appeals and parties associated with these appeals enjoyed electoral success. This changed in the aftermath of the Muzaffarnagar riots, with almost full consolidation of support behind the BJP in the 2014 election onward. Indeed, three of the major players in the pre-2013 era, BSP, SP, and RLD, contested the 2019 national elections in coalition. Even in this scenario, the BJP was ahead in 46 of 58 ACs.

When there is such large-scale Hindu-Muslim polarization, outcomes resemble what political scientist Karen Ferree calls "census elections" – where the full consolidation of Hindus takes advantage of its relative numerical dominance, even in the face of significant Muslim populations. For instance, there were 9 ACs in (undivided) Muzaffarnagar district with 41% Muslim population (figure 2). But, even here, the BJP won 8 out of 9 ACs in the 2017 state election.

In the 2022 state election, the Hindu-Muslim polarization is noticeably dampened. This region was a hotspot in the recent farmer protests, which have generated significant anger against the BJP among the Jat community. In tandem with other economic issues such as joblessness and price rise, there is some consternation against the ruling government. The core empirical question with regard to elections in this region is whether this will generate enough fragmentation in the Hindu vote to prevent a census election scenario, particularly in areas with significant Muslim populations.

3   A Detailed Look at the 2017 Election

In this section, we analyze polling booth results from the 2017 state election in Uttar Pradesh. These data have been digitized and sanitized by *Datalok* and yield records for 21,271 polling

booths across 58 ACs.

Polling booth data provide a detailed picture of local politics, as they are not the level at which elections are contested. While analysis of data at the AC level reflect strategic alliances and appeals (what is often referred to as Downsian competition), no such dynamics exist at the polling booth level. Strategic calculations by voters are a function of politics at the constituency and state level and not at the polling booth. The polling booth data, thus, provides an unvarnished picture of the consolidations at the local level for each party.

In this technical analysis, we describe how Hindu-Muslim polarization, which was still very evident in the region's 2017 electoral results, manifest at the local level. It is also an ideal election to understand how the key parties, BSP, RLD, and SP, behave as individual entities – as there was no alliance between them. In 2017, only the SP and INC were in alliance.

Two key empirical findings result from this analysis. First, polarization generated a number of "uncompetitive" polling booths in which consolidation disproportionately favored one party. In areas with presumably fewer Muslim and Jatav voters, there is a large cluster of uncompetitive polling booths that consolidated in favor of the BJP. Second, where there was party fragmentation at the local level, the relative threshold – i.e., required vote share to win a polling booth – was significantly lower. In these scenarios, parties that consolidate vote shares for one community (RLD towards the Jat community and the BSP for the Jatav community) disproportionately benefit.

3.1 Party Fragmentation and Consolidations

Throughout the rest of this article, we make use a of standard metric of party fragmentation, the effective number of parties (ENP). Intuitively, this measure provides an adjustment for the amount of vote share accrued to a party.

An example will be helpful here. Imagine a place with 4 candidates/parties standing for election. We can imagine a scenario in which the parties all have equal (25%) vote share, or we can imagine a scenario in which one party garners nearly 90% of the vote and another nearly 10% of the vote (with minimal vote for the other two parties). Clearly, all 4 parties are competitive in the first scenario, whereas really only one party has any chance of winning in the second scenario. As such, we would want the "effective" number of parties to be much higher in the first scenario as compared to the second scenario. In both scenarios, 4 parties are competing, so simply tallying the number of parties standing for election does not provide much information. The trick is to incorporate vote shares into the calculation, using a method developed by Laakso and Taagepara known as the effective number of parties index:

$$ENP = \frac{1}{\sum_{i=1}^{n} p_i^2}$$

where p_i is the proportion of votes for party i and n is the number of parties

For the scenarios described above, when all 4 parties have equal vote share the ENP is 4, and when one party has 90% vote share with another receiving 10% vote share the ENP is 1.22. This

captures the fact that in the first scenario all 4 parties are of equal competitiveness while in the second scenario is only one party that is plausibly competitive.

In order to characterize the type of competition in a polling booth with respect to ENP, we classify 4 categories: uncompetitive, bipolar, multiparty, and fragmented. Uncompetitive polling booths are those in which the ENP is less than 1.9, corresponding to a gap (margin of victory or MoV) of more than 20 percentage points between the first place and second place party in two party competition. Bipolar polling booths are those in which ENP is between 1.9 and 2.4, corresponding to a scenario in which the top two parties are within 20 percentage points of each other and a third party has at most 10% vote share. Multiparty polling booths correspond to ENP values between 2.4 and 3, where there are at least 3 parties with reasonably significant vote share. Finally, a fragmented polling booth has an ENP value of greater than 3, corresponding to a scenario in which there are at least 3 equally matched parties and often more than 3 parties with significant vote share.

(a) Effective Number of Parties
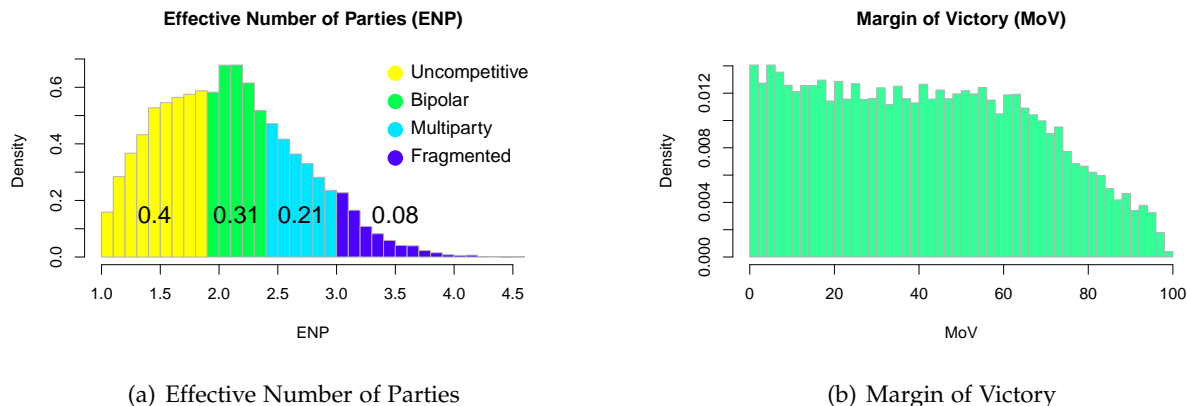
(b) Margin of Victory

Figure 4: ENP and MoV in Polling Booths in 2017 for Phase 1 ACs

Figure 4 provides an analysis of 21,271 polling booths across the 58 ACs in phase 1. Panel 4(a) provides a histogram of ENP at the polling booth level, with various colors representing whether the polling booth is uncompetitive, bipolar, multiparty, or fragmented. Panel 4(b) provides a histogram of the margin of victory (MoV) at the polling booth level.

Figure 4(a) presents a histogram of ENP values at the polling booth level categorized by type of competition. This demonstrates that a little over 40% of polling booth were uncompetitive which ostensibly saw either Hindu consolidation where there was a large majority of Hindus, or Muslim, Jat, or Jatav consolidation for the SP-INC alliance, RLD, and BSP, respectively. A little over 30% of polling booths saw some form of bipolar competition and another 30% saw either multiparty or fragmented party competition. All in all, however, polling booths were not very competitive. Only 13% of polling booth had a margin of victory less than 10 percentage points (figure 4(b)). Taken together, this analysis shows there were many local strongholds for various parties in 2017, and that the political opposition will have to wrest control of the more competitive polling booths if it is to eat into BJP's dominance in the region.

3.2 Party Performance at the Polling Booth Level

One way to understand the performance of parties at the local level is simply to look at the distribution of vote shares at the polling booth level for each major party/alliance, as has been done in figure 5. Two major lessons emerge from this figure. First, outside of the BJP, the other major parties had very few areas of significant vote share. The BSP received at least 30% vote share in just 28% of polling booths, and the SP-INC alliance reached that threshold only 23% of the time. This is consistent with the idea that these parties were limited to electoral support in areas where affiliated caste/religious identity groups were clustered. Second, there was significant polling booth level consolidation in favor of the BJP. The BJP had at least 30% vote share in 71% of polling booths and at least 50% vote share in 51% of polling booths. This suggests very significant Hindu consolidation outside of identity groups affiliated with competitor parties.
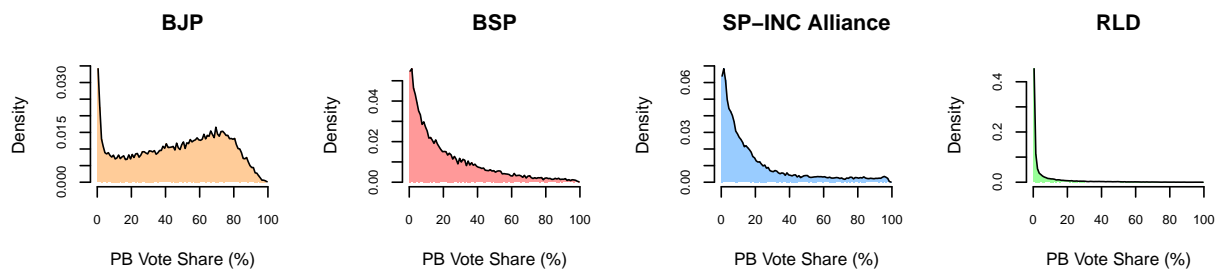

Figure 5: Vote Shares across Polling Booths in 2017 in Phase 1 ACs

These histograms display the vote shares for each major party/alliance across the 21,271 polling booths (where the party or alliance competed).

In order to test the claims we observe in the party-level distributions of vote share, we characterize party performance by type of polling booth competition (using ENP values) in figure 6.

	BJP	SP + INC	BSP	RLD
Uncompetitive	66	18	13	2
Bipolar	62	15	19	5
Multiparty	55	15	22	8
Fragmented	47	18	23	11

Figure 6: Polling Booths (%) won by Party/Alliance by Polling Booth Type in Phase 1 ACs

This table displays the percentage of polling booths won by each major party/alliance across the 4 types of polling booths classified on the basis of polling booth level ENP values. These data are calculated over the 21,271 polling booths in phase 1 ACs.

Indeed, the BJP wins the lion's share of uncompetitive polling booths (66%), and as the effective number of parties grow to make bipolar, multiparty, and fragmented polling booths, the BJP's performance weakens. In fragmented polling booths, the BJP wins 47% of polling booths – much

weaker than its performance in uncompetitive booths. By contrast, parties that are perceived to be biased towards one caste group, BSP (Jatavs) and RLD (Jats), perform much better in fragmented polling booths. This is because higher party fragmentation implies that parties must cross a lower threshold in terms of vote share to win a polling booth, something that can plausibly be done by consolidating vote among more specific caste groups.

Taken together, this analysis suggests that in order to eat into BJP's dominance, greater party fragmentation will be required at the polling booth level so that parties that appeal to a narrower base of voters can win more polling booths.

4   Concluding Thoughts

The goal of this note has been to provide a comprehensive empirical summary of recent voting trends in the constituencies heading to the polls in the first phase of the Uttar Pradesh elections, with an eye to how it maps to local social contexts. Our team is actively conducting qualitative fieldwork in these areas, and this will be the subject of another note.

A few broad observations will be helpful to understand how the data analysis presented here can shed light on what is taking place today in UP. After the recent farmer protests and the coronavirus pandemic, there is little doubt that economic issues like joblessness and price rise are heard in this region much more often as compared to 5 years ago, and that Hindu-Muslim polarization is less visible. The enmity between certain communities, most notably between Jats and Muslims, has faded considerably and the current alliance between the SP and RLD is working reasonably well. Jatavs, with many expressing grievances against the current government and the SP in addition to an identity-based link to BSP, do not seem to be adding significantly to the tally of either the BJP or the SP-RLD alliance this time around.

However, as this note shows in some detail, this arithmetic will be insufficient to displace the BJP in the region, with significant consolidation among other Hindu groups in 2017. If the SP-RLD is to make a major dent in BJP's electoral fortunes it will need more than its identity-based alliances. It will require greater fragmentation among Hindus. The depth of economic anxieties may be a deciding factor here.