Supplementary Materials

December 1, 2019

Contents

A1 RGGVY Targeting: Additional Material

A1.1 Control Variables and Split Samples

Because SC population percentages could be correlated with confounding variables, we control for a select set of variables that could influence both the SC population percentage and the probability of RGGVY implementation. For each control variable, we also explain why it is not a "bad control" in that it would be influenced by SC population (e.g., Angrist and Pischke, 2009). For summary statistics, see Table A2.

We begin with the logarithmized distance between the village and the closest town. Because all towns in Uttar Pradesh are now electrified, the cost and ease of implementing rural electrification works in nearby villages is much lower than the cost of such works in faraway villages. At the same time, the social bias against Dalits in Uttar Pradesh may mean that they tend to live farther away from towns.

Next, we control for earlier village-level electrification status, as per the 2001 Census of India (unfortunately, this earlier census does not contain household electrification percentages by village). Obviously, the status of village electrification in 2001 is a strong predictor for the need for RGGVY. At the same time, the geographic distribution of village electrification turns out be related to SC percentage in the village population (see below for details).

We also control for the logarithmized population of the village. Larger villages tend to have higher electrification rates to begin with, so they may not need the RGGVY. At the same time, larger villages tend to have more diverse populations, and thus their SC shares are much less likely to be zero than those of smaller villages.

Because RGGVY implementation requires infrastructure, we also control for the presence of a paved road. Given that Dalits historically tend to live in more remote and poorly connected villages, the presence of a paved road is also correlated with the SC population.

We include electoral constituency fixed effects in some models. These fixed effects allow us to compare villages close to each other and sharing similar political histories. It also helps us rule out competing explanations for our findings, such as those based on Uttar Pradesh's location with

Variable	Description	Source
RGGVY	Implementation of RGGVY (= 1) between April 2005 and Oct. 2014	Rural Electrification Corporation of India
Electricity	Percentage of households with grid electricity access as of 2011 (and 2001)	Census 2011
Domestic Electricity (2001)	Village is electrified as of 2001 (= 1)	Census 2001
Share SC	Share of a village's population who belongs to SC (or ST) as of 2011 (and 2001)	Census 2011 (and 2001)
Pucca Road	Indicator denoting the presence of a pucca road (= 1)	Census 2011
Distance (log)	Log distance between the village and the closest towns	Census 2011
Population (log)	Log population of the village	Census 2011
Literacy Rate (%)	Village literacy rate (%)	Census 2001
# Coop Commercial Banks	Number of cooperative banks	Census 2001
Irrigated Land (log)	Log area of irrigated land	Census 2001
Mean Light	Average nighttime luminosity, 1995-2004	NOAA satellite data
BSP Win	BSP won this constituency (= 1)	Election Commission of India
BSP Margin	Margin of victory/loss for BSP (= 1)	Election Commission of India
Caste background of MLA	Whether an MLA is SC or not	Authors' own data

Table A1: Data sources.

respect to national electricity sources.

A1.2 Summary Statistics

- Table A2 shows the summary statistics for the full sample at the village level.
- Tables A3 and A3 show the summary statistics for the full sample by districts for the presence of SCs.
- Tables A5 and A5 show the summary statistics for the full sample by districts for the implementation of RGGVY.

Summary Statistic					
	Mean	S.D.	Min.	Max	Obs.
RGGVY	31.08	46.28	0	100	96557
Domestic Electricity (2001)	35.84	47.95	0	100	96557
Lighting Source: Electricity	23.41	24.02	0	100	96557
Share SC (%)	24.57	20.70	0	100	96557
BSP Margin	-0.83	8.62	-49	22	52833
BSP Win	0.26	0.44	0	1	96557
Population (log)	6.91	1.10	1	11	96557
Distance (log)	2.43	1.10	0	5	90683
Pucca Road	0.66	0.47	0	1	96196
Lack of Asset	10.94	10.24	0	100	96557
Literacy Rate (%)	55.75	11.19	0	100	96557

Table A2: Summary statistics for the entire sample. The unit of analysis is a village.

Figure A1: Geographic distribution of SC share by village.

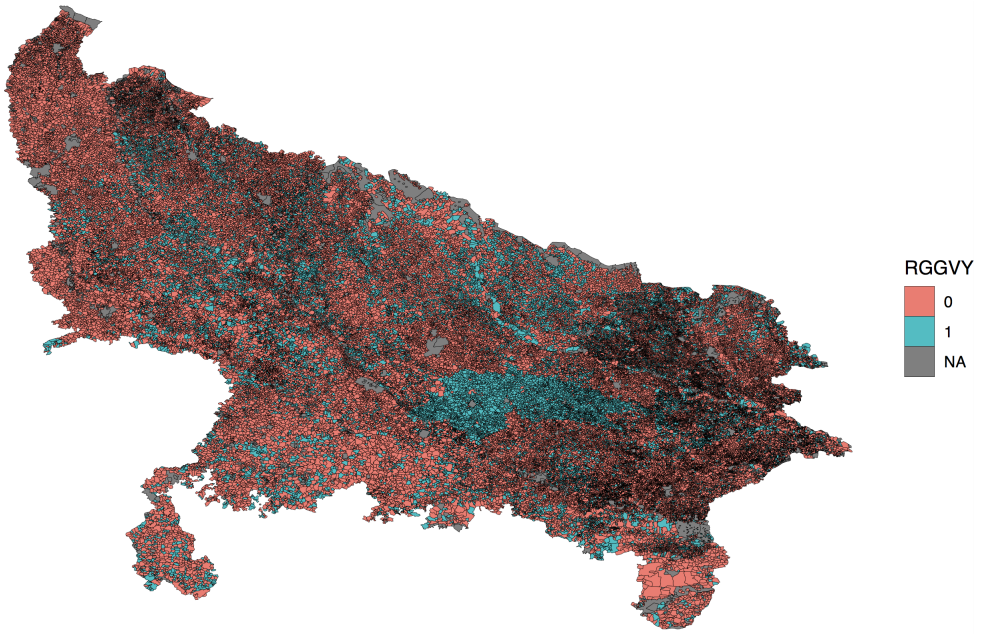

Figure A2: Geographic distribution of RGGVY implementation at the village level.

Summary Statistic: SC/ST by District (part A)

	mean	p25	p50	p75	min	max
Agra	21.6	9.0	19.0	30.4	0.0	98.7
Aligarh	23.8	11.4	21.1	32.5	0.0	100.0
Allahabad	25.3	12.1	23.0	34.6	0.0	100.0
Ambedkar Nagar	28.4	16.2	26.8	38.1	0.0	100.0
Auraiya	30.7	15.7	28.0	42.5	0.0	100.0
Azamgarh	27.5	7.2	23.2	39.6	0.0	100.0
Baghpat	13.1	6.4	11.7	16.9	0.0	73.3
Bahraich	16.4	6.9	12.9	22.3	0.0	97.2
Ballia	18.4	0.9	14.0	27.0	0.0	100.0
Balrampur	16.2	7.7	13.2	20.3	0.0	100.0
Banda	22.2	11.3	20.3	31.2	0.0	99.3
Barabanki	30.6	17.4	28.9	41.2	0.0	100.0
Bareilly	16.3	3.0	11.6	23.3	0.0	100.0
Basti	21.8	8.5	19.5	30.8	0.0	100.0
Bijnor	26.6	2.9	22.2	41.9	0.0	100.0
Budaun	18.3	3.0	13.2	26.2	0.0	100.0
Bulandshahar	23.6	9.8	20.4	33.1	0.0	100.0
Chandauli	28.0	6.9	23.9	40.3	0.0	100.0
Chitrakoot	28.6	14.0	25.3	37.1	0.0	100.0
Deoria	20.4	8.4	17.7	27.5	0.0	100.0
Etah	17.9	3.8	13.5	26.0	0.0	100.0
Etawah	26.5	11.6	23.1	38.0	0.0	100.0
Faizabad	24.5	14.7	22.8	32.1	0.0	99.9
Farrukhabad	17.1	5.2	13.1	24.2	0.0	100.0
Fatehpur	27.3	15.6	25.8	37.1	0.0	100.0
Firozabad	19.7	4.3	15.9	28.5	0.0	100.0
Gautam Buddha Nagar	20.2	8.0	17.6	26.9	0.0	95.4
Ghaziabad	21.1	8.0	18.0	30.2	0.0	100.0
Ghazipur	22.4	0.0	16.8	33.4	0.0	100.0
Gonda	17.5	8.4	15.2	23.3	0.0	100.0
Gorakhpur	25.4	8.0	21.0	35.9	0.0	100.0
Hamirpur	22.4	13.5	22.1	30.0	0.0	100.0
Hardoi	33.0	14.7	29.6	47.9	0.0	100.0
Hathras	27.6	12.7	24.4	37.2	0.0	100.0

Table A3: Summary statistics on the presence of SC, by district.

Summary Statistic: SC/ST by District (part B)

	mean	p25	p50	p75	min	max
Jalaun	29.2	15.3	28.1	39.6	0.0	100.0
Jaunpur	23.6	6.7	20.4	33.6	0.0	100.0
Jhansi	31.8	21.7	31.6	41.3	0.0	100.0
Jyotiba Phule Nagar	18.9	1.3	12.4	29.1	0.0	100.0
Kannauj	21.0	9.9	18.0	28.2	0.0	100.0
Kanpur Dehat	26.6	13.7	25.4	36.8	0.0	100.0
Kanpur Nagar	28.8	16.4	26.9	38.5	0.0	100.0
Kaushambi	36.8	24.2	35.8	48.3	0.0	100.0
Kheri	33.4	16.8	29.6	44.3	0.0	100.0
Kushinagar	18.6	9.2	16.2	25.3	0.0	100.0
Lalitpur	27.2	16.9	26.3	36.0	0.0	100.0
Lucknow	41.3	27.7	40.3	53.3	0.0	100.0
Mahoba	27.1	18.0	25.8	35.1	0.0	100.0
Mahrajganj	20.1	10.9	18.3	26.8	0.0	96.5
Mainpuri	20.4	6.7	17.4	28.2	0.0	100.0
Mathura	21.1	9.9	17.9	28.7	0.0	100.0
Mau	26.8	2.5	20.6	40.3	0.0	100.0
Meerut	22.4	7.9	19.9	32.4	0.0	100.0
Mirzapur	29.4	9.1	25.3	45.7	0.0	100.0
Moradabad	19.7	2.6	13.8	29.5	0.0	100.0
Muzaffarnagar	18.5	6.1	14.5	24.7	0.0	100.0
Pilibhit	17.0	2.4	12.8	23.8	0.0	100.0
Pratapgarh	22.7	12.4	21.2	30.7	0.0	100.0
Rae Bareli	33.0	22.2	31.7	42.8	0.0	100.0
Rampur	16.6	1.2	8.6	25.3	0.0	100.0
Saharanpur	28.2	10.1	25.4	40.4	0.0	100.0
Sant Kabir Nagar	23.2	8.4	19.5	32.4	0.0	100.0
Sant Ravidas Nagar Bhadohi	22.7	3.5	19.2	33.1	0.0	100.0
Shahjahanpur	20.4	4.4	14.3	28.7	0.0	100.0
Shrawasti	19.6	10.7	17.1	25.3	0.0	99.4
Siddharthnagar	17.8	6.5	14.7	25.0	0.0	100.0
Sitapur	37.8	23.1	36.3	50.1	0.0	100.0
Sonbhadra	43.7	22.0	40.8	64.2	0.0	100.0
Sultanpur	23.8	13.1	22.0	31.8	0.0	100.0
Unnao	36.0	22.3	33.7	47.4	0.0	100.0
Varanasi	14.8	3.6	14.9	27.4	0.0	100.0
Total	24.6	9.0	21.0	34.9	0.0	100.0

Table A4: Summary statistics on the presence of SC, by district.

Summary Statistic: RGGVY by District (Part A)

	Mean	25th pctl	50th pctl	75th pctl	Min	Max
Agra	15.6	0.0	0.0	0.0	0.0	100.0
Aligarh	28.7	0.0	0.0	100.0	0.0	100.0
Allahabad	35.3	0.0	0.0	100.0	0.0	100.0
Ambedkar Nagar	6.8	0.0	0.0	0.0	0.0	100.0
Auraiya	37.9	0.0	0.0	100.0	0.0	100.0
Azamgarh	49.0	0.0	0.0	100.0	0.0	100.0
Baghpat	0.0	0.0	0.0	0.0	0.0	0.0
Bahraich	46.6	0.0	0.0	100.0	0.0	100.0
Ballia	27.3	0.0	0.0	100.0	0.0	100.0
Balrampur	24.1	0.0	0.0	0.0	0.0	100.0
Banda	22.6	0.0	0.0	0.0	0.0	100.0
Barabanki	27.8	0.0	0.0	100.0	0.0	100.0
Bareilly	25.6	0.0	0.0	100.0	0.0	100.0
Basti	36.0	0.0	0.0	100.0	0.0	100.0
Bijnor	13.9	0.0	0.0	0.0	0.0	100.0
Budaun	30.8	0.0	0.0	100.0	0.0	100.0
Bulandshahar	18.4	0.0	0.0	0.0	0.0	100.0
Chandauli	20.5	0.0	0.0	0.0	0.0	100.0
Chitrakoot	34.7	0.0	0.0	100.0	0.0	100.0
Deoria	12.1	0.0	0.0	0.0	0.0	100.0
Etah	48.7	0.0	0.0	100.0	0.0	100.0
Etawah	29.6	0.0	0.0	100.0	0.0	100.0
Faizabad	34.4	0.0	0.0	100.0	0.0	100.0
Farrukhabad	35.7	0.0	0.0	100.0	0.0	100.0
Fatehpur	33.8	0.0	0.0	100.0	0.0	100.0
Firozabad	33.7	0.0	0.0	100.0	0.0	100.0
Gautam Buddha Nagar	16.3	0.0	0.0	0.0	0.0	100.0
Ghaziabad	1.4	0.0	0.0	0.0	0.0	100.0
Ghazipur	9.4	0.0	0.0	0.0	0.0	100.0
Gonda	53.0	0.0	100.0	100.0	0.0	100.0
Gorakhpur	18.3	0.0	0.0	0.0	0.0	100.0
Hamirpur	28.0	0.0	0.0	100.0	0.0	100.0
Hardoi	40.7	0.0	0.0	100.0	0.0	100.0
Hathras	17.4	0.0	0.0	0.0	0.0	100.0

Table A5: Summary statistics on the implementation of RGGVY, by district.

Summary Statistic: RGGVY by District (Part B)

	Mean	25th pctl	50th pctl	75th pctl	Min	Max
Jalaun	11.4	0.0	0.0	0.0	0.0	100.0
Jaunpur	21.1	0.0	0.0	0.0	0.0	100.0
Jhansi	24.5	0.0	0.0	0.0	0.0	100.0
Jyotiba Phule Nagar	53.5	0.0	100.0	100.0	0.0	100.0
Kannauj	35.6	0.0	0.0	100.0	0.0	100.0
Kanpur Dehat	34.6	0.0	0.0	100.0	0.0	100.0
Kanpur Nagar	17.4	0.0	0.0	0.0	0.0	100.0
Kaushambi	25.6	0.0	0.0	100.0	0.0	100.0
Kheri	36.9	0.0	0.0	100.0	0.0	100.0
Kushinagar	28.0	0.0	0.0	100.0	0.0	100.0
Lalitpur	37.3	0.0	0.0	100.0	0.0	100.0
Lucknow	4.0	0.0	0.0	0.0	0.0	100.0
Mahoba	42.3	0.0	0.0	100.0	0.0	100.0
Mahrajganj	24.6	0.0	0.0	0.0	0.0	100.0
Mainpuri	33.8	0.0	0.0	100.0	0.0	100.0
Mathura	4.1	0.0	0.0	0.0	0.0	100.0
Mau	45.1	0.0	0.0	100.0	0.0	100.0
Meerut	0.0	0.0	0.0	0.0	0.0	0.0
Mirzapur	32.3	0.0	0.0	100.0	0.0	100.0
Moradabad	20.7	0.0	0.0	0.0	0.0	100.0
Muzaffarnagar	0.0	0.0	0.0	0.0	0.0	0.0
Pilibhit	34.2	0.0	0.0	100.0	0.0	100.0
Pratapgarh	19.4	0.0	0.0	0.0	0.0	100.0
Rae Bareli	97.2	100.0	100.0	100.0	0.0	100.0
Rampur	33.5	0.0	0.0	100.0	0.0	100.0
Saharanpur	0.0	0.0	0.0	0.0	0.0	0.0
Sant Kabir Nagar	36.9	0.0	0.0	100.0	0.0	100.0
Sant Ravidas Nagar Bhadohi	12.4	0.0	0.0	0.0	0.0	100.0
Shahjahanpur	35.8	0.0	0.0	100.0	0.0	100.0
Shrawasti	46.3	0.0	0.0	100.0	0.0	100.0
Siddharthnagar	45.2	0.0	0.0	100.0	0.0	100.0
Sitapur	39.6	0.0	0.0	100.0	0.0	100.0
Sonbhadra	24.6	0.0	0.0	0.0	0.0	100.0
Sultanpur	90.0	100.0	100.0	100.0	0.0	100.0
Unnao	36.6	0.0	0.0	100.0	0.0	100.0
Varanasi	0.0	0.0	0.0	0.0	0.0	0.0
Total	31.1	0.0	0.0	100.0	0.0	100.0

Table A6: Summary statistics on the implementation of RGGVY, by district.

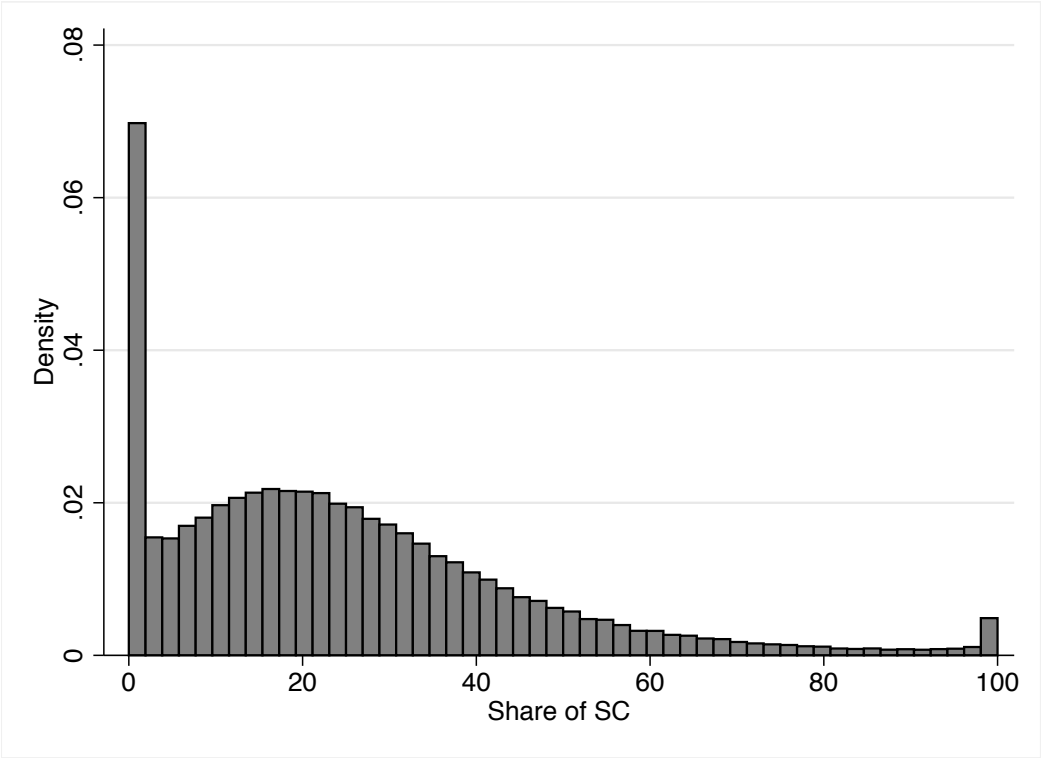

Figure A3: Distribution of the share of scheduled caste members per village (entire sample).

Figure A3 demonstrates that there is considerable variation across villages in SC population. The x-axis shows the SC population percentage on a 0-100 scale, and the y-axis shows the density of different percentages. While a large number of villages have no SC population at all and there are also villages with only SC people, the vast majority of the villages fall on a right-skewed normal distribution. The average SC percentage in our dataset is 24.6%.

A2 Pre-RGGVY Rural Electrification

In conducting our study, we consider the relationship between village electrification and SC population *before* the RGGVY begins. To achieve this goal, we use the 2001 Census of India. While this earlier census unfortunately does not contain information about household electrification, it does allow us to compute SC population percentages and assess village electrification. Table A7 regresses the electrification status in 2001 on the village SC population percentage. The SC percentage is actually *positively* correlated with the likelihood of village electrification. Increasing the

SC percentage by 10 points, for example, increases the probability of village electrification by approximately 1 percentage point across the models – an association that is sensitive neither to the inclusion of fixed effects nor to that of control variables.

If the Dalit population is generally underprivileged in India and Uttar Pradesh, why would their villages enjoy higher levels of electrification before the RGGVY? To understand the initially puzzling relationship between SC population percentage and village electrification, Table A8 offers summary statistics by the decile of SC population percentage. As the table shows, the surprising result is almost entirely driven by villages with no Dalits at all: the difference between the 2nd and 10th decile in the probability of village electrification is only 6 percentage points, while the difference between the 1st and 2nd decile alone is 7 percentage points. Because villages without Dalits tend to be very small (average population: 429), it is unsurprising that they have no village electrification. Villages in all other deciles are larger, so they have higher probabilities of village electrification and road construction, but their development outcomes (no assets, literacy) are not very different.

- Table A7 regresses the electrification status in 2001 on the village SC population percentage. The dependent variable in all models is electrification status in 2001, which is 100 if the village is electrified and zero otherwise. Models 2-6 include constituency fixed effects, and standard errors are clustered by constituency throughout.

- Table A8 offers summary statistics by the decile of SC population percentage.

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	0.11***	0.10***	0.10***	0.10***	0.10***	0.10***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance (log)			-0.98***			-0.90***
			(0.21)			(0.21)
Population (log)				4.07***		3.92***
				(0.31)		(0.32)
Pucca Road					3.20***	1.32***
					(0.48)	(0.47)
Constant	33.15***					
	(0.77)					
Constituency FE		✓	✓	✓	✓	✓
Observations	96557	96557	90683	96557	96196	90454
R^2	0.00	0.00	0.00	0.01	0.00	0.01
# Clusters	402	402	401	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Dependent variable: electrification status in 2001 (= 100 if electrified; 0 otherwise). All models estimated with constituency fixed effects. The standard errors are clustered by constituency.

Decile	Share of SC (cutoff)	Dom. Electricity (2001)	Population	Road	No assets	Literacy	Distance to nearest town
1st	0%	26.4%	429	50.3%	9.1%	57%	15.6
2nd	6.2%	33.2%	1460	64.2%	11.2%	53.8%	16.3
3rd	11.6%	35.5%	1947	67.0%	10.8%	54.6%	17.1
4th	16.3%	36.7%	2121	68.3%	10.6%	55.5%	17.6
5th	21.0%	37.0%	2080	69.3%	10.4%	56.4%	17.3
6th	25.8%	36.4%	2011	70.0%	10.6%	56.6%	17.4
7th	31.5%	37.1%	1914	69.5%	10.5%	57.1%	17.5
8th	38.9%	38.4%	1715	70.2%	10.9%	56.9%	17.6
9th	51.3%	38.8%	1450	68.8%	11.3%	56.2%	17.8
10th	Above 51.3%	39.2%	976	66.8%	14.0%	53.2%	17.6

Table A8: Dependent variable: village electrification rate/probability (%) in 2001 by decile of SC population percentage in a given village. The second decile is somewhat smaller than the others because villages with zero percent SC are excluded from this group.

A3 Regression Discontinuity: Identifying Assumptions

In an RDD analysis, local average treatment effects are identified by quantifying a discontinuous jump in the outcome at the threshold (Imbens and Lemieux, 2008). In our case, this means comparing RGGVY implementation between electoral constituencies that were barely won or lost by the BSP. The basic identifying assumption is that while the outcome may be related to the forcing variable, such as the margin of victory, the sharp discontinuity at the cut-off – in our case, BSP victory – allows the estimation of local average treatment effects for villages within electoral constituencies in which the BSP barely won or lost.

The identifying assumption can be tested in several ways. The first is to compare pre-treatment covariate values in constituencies barely won or lost by the BSP. These balance statistics are provided in Table A12 to A14. As the table shows, the treatment (BSP victory by a narrow margin) and control (BSP loss by a narrow margin) are statistically indistinguishable for pre-treatment covariates.

Following McCrary (2008), we also examine any discontinuities at the cut-off (Figure A6). The test shows that there is no suspicious discontinuity, alleviating concerns about electoral fraud and other irregularities in the conduct of election.

To scrutinize the external validity of the results, we also replicate them in the full sample. While the full sample estimation does not admit causal inference, it can be used to see whether the correlations in the data are broadly consistent with the results from the close elections. If they are consistent, this observation alleviates concerns about close elections being a special case without external validity.

A4 Regression Discontinuity: Summary Statistics

- Figure A4 shows the margin of victory for each constituency-election in the RDD sample (5% margin of victory).

- Table A9 summarizes the RDD sample. The upper panel summarizes the data at the village level; the lower panel summarizes the data at the constituency-election level. In total, we

have 235 close constituency-elections when the sample is restricted to a 5% margin of victory.

- Table A10 compares BSP and non-BSP MLAs. As the table shows, both the candidate and constituency characteristics are mostly similar. The only exception – an unsurprising one – is that BSP MLAs tend to come from SC-reserved constituencies.
- Figure A5 shows the kernel density function for the SC share in the RDD sample.
- Table A5 is the histogram of BSP wins and losses (i.e. when it came second) based on a $+/-5$ percent margin.
- Table A11 reports the summary statistics for the main variables used for the regression discontinuity analysis (using a $+/-5$ percent margin).

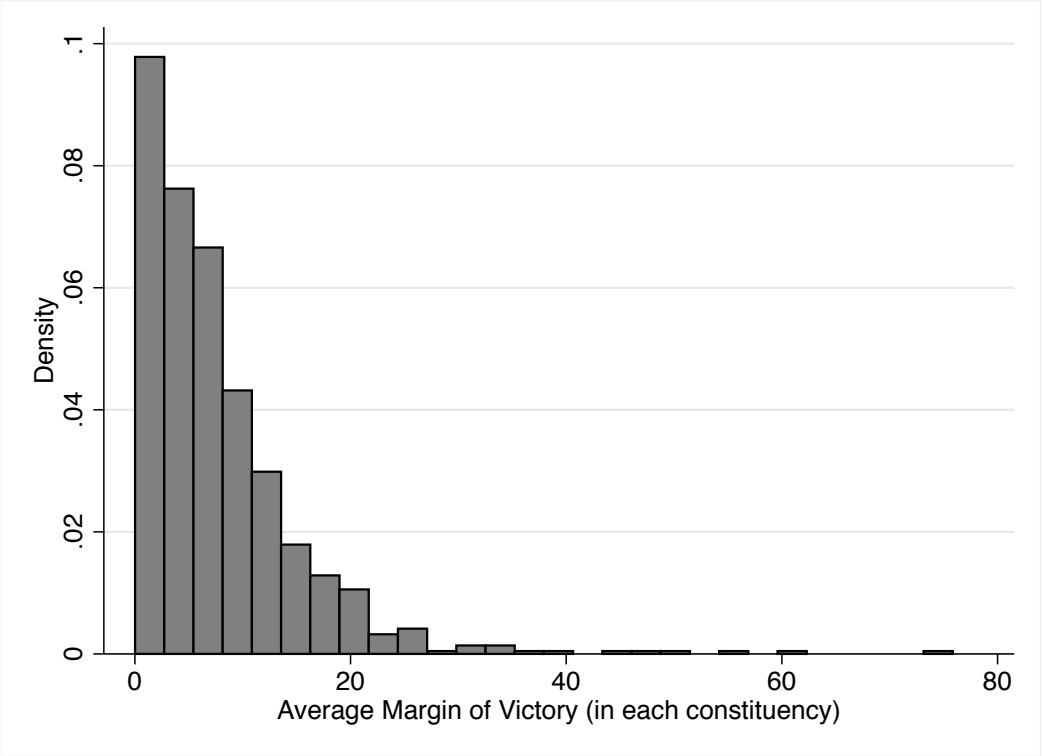

Figure A4: Distribution of the margin of victory for the winning party against the first runner-up in each constituency-election.

Village Level

Category	*Won by BSP*	*Lost by BSP*	*Total # of villages*
All village-elections	77,617	115,497	193,114
All village-elections, BSP top-2	77,617	53,723	131,340
All villages, 2002, BSP top-2	25,556	27,277	52,833
All villages, 2007, BSP top-2	52,061	26,446	78,507
Village-elections, BSP top-2, 1% win/loss margin	7,170	6,916	14,086
Village-elections, BSP top-2, 2% win/loss margin	14,266	12,527	26,793
Village-elections, BSP top-2, 5% win/loss margin	34,281	27,798	62,079

Constituency Level

Category	*Won by BSP*	*Lost by BSP*	*Total # of constituencies*
All constituency-elections	303	501	804
All constituency-elections, BSP top-2	303	217	520
All constituencies, 2002, BSP top-2	98	108	206
All constituencies, 2007, BSP top-2	205	109	314
Constituency-elections, BSP top-2, 1% win/loss margin	27	29	56
Constituency-elections, BSP top-2, 2% win/loss margin	51	50	101
Constituency-elections, BSP top-2, 5% win/loss margin	129	106	235

Table A9: Summary of the RDD sample.

	Full Sample		5% Sample	
	Non-BSP	BSP	Non-BSP	BSP
Candidate Characteristics				
Male	0.934	0.951	0.96	0.974
Higher Education	0.67	0.568	0.65	0.623
Number of Criminal Charges	1.107	1.01	1.25	0.961
Asset (10,000 rupees)	1,448.601	832.718	1,094.880	649.169
Debt (10,000 rupees)	100.952	136.879	76.930	87.660
Constituency Characteristics				
SC Constituency	0.142	0.301	0.11	0.208
Domestic Electricity (2001)	0.342	0.32	0.318	0.308
Household Electrification Rate (2011)	0.267	0.248	0.256	0.229
Literacy Rate (2001)	0.558	0.563	0.556	0.56
Literacy Rate (2011)	0.424	0.432	0.421	0.429
Number of Electors (10,000 people)	28.422	27.941	28.502	28.258
Num. Obs.	197	206	100	77

Table A10: Comparison between BSP and non-BSP MLAs.

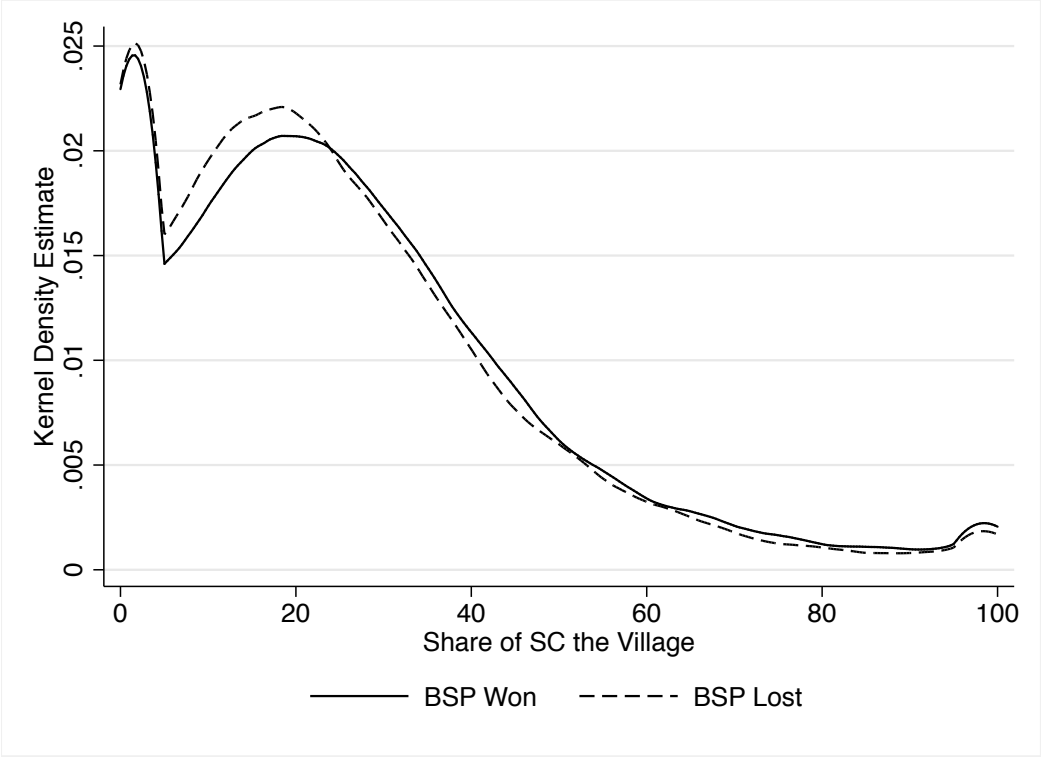

Figure A5: Histogram (kernel density function) of the share of SC in the sample, split by cases where BSP won and BSP came second, when the margin of victory is $+/-5$ percent.

Summary Statistic

	Mean	S.D.	Min.	Max	Obs.
RGGVY	30.40	46.00	0	100	62079
Domestic Electricity (2001)	34.82	47.64	0	100	62079
Lighting Source: Electricity	21.84	23.10	0	100	62079
Share SC (%)	25.30	21.22	0	100	62079
BSP Margin	0.27	2.77	-5	5	62079
BSP Win	0.55	0.50	0	1	62079
Population (log)	6.85	1.11	1	11	62079
Distance (log)	2.44	1.09	0	5	58571
Pucca Road	0.64	0.48	0	1	61800
Lack of Asset	10.84	10.22	0	100	62079
Literacy Rate (%)	56.11	11.09	0	100	62079

Table A11: Summary statistics for the sample used in the RDD study (observations with a margin below 5%). The unit of analysis is village-election.

A5 Regression Discontinuity: Balance Statistics and Density Tests

- Tables A12-A14 show the balance statistics for the 1%, 2%, and 5% RDD samples.
- Figure A6 shows the results of a McCrary (2008) density test.

Balance Statistic							
	BSP=0			BSP=1			P-value
	Mean	S.D.	Obs.	Mean	S.D.	Obs.	of Difference
Margin of Victory	0.49	0.29	29	0.46	0.26	27	0.67
Total Scheduled Castes Population of Village	416.01	153.25	29	375.72	134.61	27	0.30
Total Scheduled Tribes Population of Village	6.60	18.14	29	4.89	11.42	27	0.68
Total Population of Village	1718.45	596.54	29	1633.89	465.20	27	0.57
Area of Village (hectares)	253.77	165.12	29	186.54	89.90	27	0.07
Number of Co-operative Commercial Bank	0.02	0.01	29	0.02	0.02	27	0.95
Credit Societies (Y/N)	0.07	0.04	29	0.09	0.04	27	0.15
Paved Road	0.59	0.13	29	0.60	0.13	27	0.77
Distance from the Nearest Town (km)	11.10	3.56	29	8.76	1.94	27	<0.01**
Power Supply (A/NA)	0.68	0.17	29	0.75	0.18	27	0.18
Domestic Electricity (2001)	0.36	0.13	29	0.31	0.12	27	0.10
Agricultural Electricity (2001)	0.23	0.13	29	0.23	0.12	27	0.91
Electricity (other purposes) (2001)	0.02	0.02	29	0.02	0.02	27	0.21
Electricity (all purposes) (2001)	0.25	0.19	29	0.37	0.23	27	0.05**
Total Irrigated Area	186.39	147.11	29	206.10	474.68	27	0.83
Unirrigated Area	88.30	136.55	29	41.85	53.19	27	0.10

Table A12: Balance statistic at the constituency-election level. Village-elections where the winning margin was below 1 percent, and where neither of the top-2 candidates were members of BSP, were dropped. The summary statistics of each variable were then computed by constituency-election. The p-value is based on a t test where the null hypothesis that the means are equal. *=p<0.05, **=p<0.01.

Balance Statistic							
	BSP=0			BSP=1			P-value
	Mean	S.D.	Obs.	Mean	S.D.	Obs.	of Difference
Margin of Victory	0.94	0.61	50	0.95	0.60	51	0.96
Total Scheduled Castes Population of Village	415.83	163.59	50	407.57	170.79	51	0.80
Total Scheduled Tribes Population of Village	7.74	17.80	50	3.24	8.51	51	0.11
Total Population of Village	1773.85	670.28	50	1697.81	688.36	51	0.58
Area of Village (hectares)	248.56	167.96	50	222.03	164.19	51	0.42
Number of Co-operative Commercial Bank	0.02	0.02	50	0.02	0.02	51	0.76
Credit Societies (Y/N)	0.07	0.04	50	0.08	0.04	51	0.22
Paved Road	0.60	0.15	50	0.62	0.15	51	0.46
Distance from the Nearest Town (km)	10.37	3.52	50	9.30	2.08	51	0.07
Power Supply (A/NA)	0.68	0.18	50	0.75	0.18	51	0.07
Domestic Electricity (2001)	0.37	0.14	50	0.32	0.12	51	0.05*
Agricultural Electricity (2001)	0.21	0.13	50	0.23	0.10	51	0.48
Electricity (other purposes) (2001)	0.02	0.02	50	0.02	0.02	51	0.57
Electricity (all purposes) (2001)	0.25	0.18	50	0.36	0.23	51	0.01**
Total Irrigated Area	158.13	122.51	50	171.55	348.69	51	0.80
Unirrigated Area	68.78	112.23	50	51.43	86.38	51	0.39

Table A13: Balance statistic at the constituency-election level. Village-elections where the winning margin was below 2 percent, and where neither of the top-2 candidates were members of BSP, were dropped. The summary statistics of each variable were then computed by constituency-election. The p-value is based on a t test where the null hypothesis that the means are equal. *=p<0.05, **=p<0.01.

Balance Statistic							
	BSP=0			BSP=1			P-value
	Mean	S.D.	Obs.	Mean	S.D.	Obs.	of Difference
Margin of Victory	2.26	1.50	106	2.41	1.41	129	0.44
Total Scheduled Castes Population of Village	399.95	167.34	106	416.86	176.11	129	0.45
Total Scheduled Tribes Population of Village	6.86	16.02	106	13.46	79.49	129	0.40
Total Population of Village	1784.53	632.58	106	1744.80	725.18	129	0.66
Area of Village (hectares)	236.47	146.58	106	246.75	197.21	129	0.66
Number of Co-operative Commercial Bank	0.02	0.02	106	0.02	0.01	129	0.52
Credit Societies (Y/N)	0.08	0.04	106	0.08	0.04	129	0.84
Paved Road	0.61	0.15	106	0.60	0.13	129	0.75
Distance from the Nearest Town (km)	9.95	3.58	106	9.75	2.81	129	0.63
Power Supply (A/NA)	0.71	0.18	106	0.71	0.18	129	0.93
Domestic Electricity (2001)	0.37	0.14	106	0.34	0.12	129	0.15
Agricultural Electricity (2001)	0.23	0.13	106	0.25	0.13	129	0.33
Electricity (other purposes) (2001)	0.02	0.02	106	0.02	0.02	129	0.33
Electricity (all purposes) (2001)	0.29	0.20	106	0.30	0.20	129	0.58
Total Irrigated Area	176.48	264.17	106	151.78	239.28	129	0.45
Unirrigated Area	69.04	162.54	106	67.29	178.48	129	0.93

Table A14: Balance statistic at the constituency-election level. Village-elections where the winning margin was below 5 percent, and where neither of the top-2 candidates were members of BSP, were dropped. The summary statistics of each variable were then computed by constituency-election. The p-value is based on a t test where the null hypothesis that the means are equal. *=p<0.05, **=p<0.01.

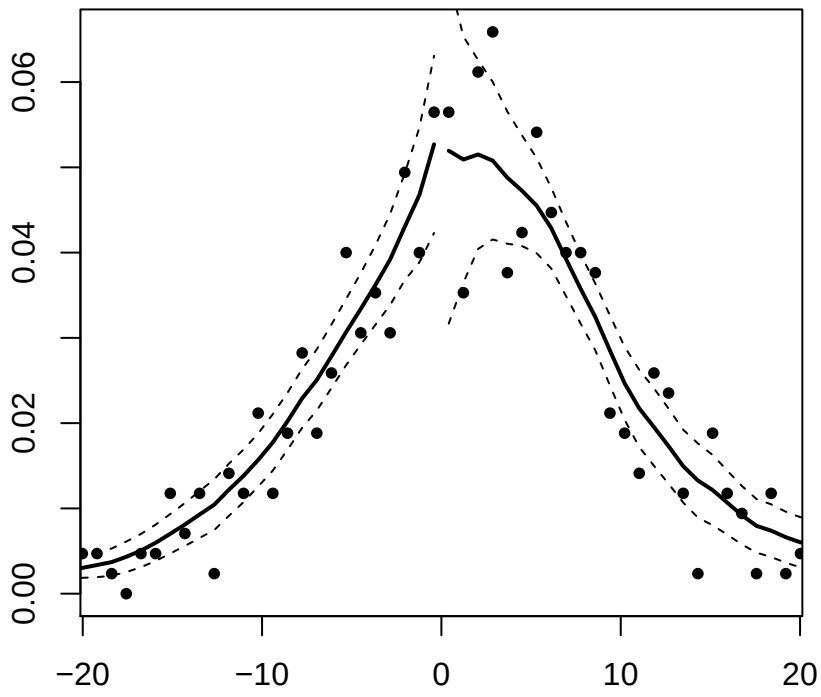

Figure A6: McCrary (2008) density test shows that there is no suspicious discontinuity in the treatment assignment around the cutoff. The p-value for rejecting the null hypothesis is 0.77.

A6 Regression Discontinuity: Additional Analysis

- In Table A15, we estimate the correlation between a BSP victory and RGGVY implementation in the full sample.
- Table A16 reports the estimates of the RDD analysis, limiting the sample to 2002.
- Table A17 reports the estimates of the RDD analysis, limiting the sample to 2007.
- Table A18 reports the estimates of the RDD analysis, but adds an interaction effect between the treatment (a BSP win) and the share of SC in the village. The sample is limited to 2002.
- Table A19 reports the estimates of the RDD analysis, but adds an interaction effect between the treatment (a BSP win) and the share of SC in the village. The sample is limited to 2007.
- Figures A7-A9 reports the regression discontinuity graph. Unlike traditional RDD figures, we bin observations to account for the dichotomous nature of the dependent variable (RGGVY).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BSP Win	-3.06*	-3.32*	-0.53	-0.47	-2.07	-2.29	0.02	0.04
	(1.78)	(1.93)	(3.13)	(3.13)	(1.93)	(2.08)	(3.40)	(3.39)
2007 Election		0.91*	0.72	0.72		0.74	0.46	0.47
		(0.54)	(1.15)	(1.15)		(0.54)	(1.15)	(1.14)
BSP Margin			-0.09	-0.04			-0.05	-0.02
			(0.19)	(0.26)			(0.19)	(0.25)
BSP Win * Margin				-0.10				-0.08
				(0.32)				(0.32)
BSP Win * Share SC					-0.02	-0.02	-0.02	-0.02
					(0.03)	(0.03)	(0.03)	(0.03)
Share SC (%)					-0.16***	-0.16***	-0.16***	-0.16***
					(0.03)	(0.03)	(0.03)	(0.03)
Constant	32.31***	31.96***	29.88***	30.18***	36.15***	35.86***	34.08***	34.32***
	(1.52)	(1.40)	(2.20)	(2.42)	(1.59)	(1.49)	(2.48)	(2.74)
Observations	193114	193114	131340	131340	193114	193114	131340	131340
R^2	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
# Clusters	402	402	340	340	402	402	340	340

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A15: Full sample, mimicking both the RDD and the RDD with interactions. Dependent variable: RGGVY (if present, RGGVY= 100). The standard errors are clustered by constituency.

	Margin<1%			Margin<2%			Margin<5%		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BSP Win	6.16	29.42**	32.68**	0.92	13.57	14.04	0.35	1.93	2.06
	(6.98)	(13.70)	(14.20)	(7.35)	(10.27)	(10.24)	(4.90)	(8.57)	(8.43)
BSP Margin		-27.36**	-11.59		-8.07	-3.98		-0.33	0.99
		(11.21)	(6.88)		(6.24)	(8.31)		(1.64)	(2.14)
BSP Win * Margin			-37.24			-8.85			-2.78
			(23.83)			(11.76)			(3.19)
Constant	27.70***	16.69***	23.04***	32.06***	25.70***	28.93***	30.01***	29.20***	32.47***
	(2.62)	(5.25)	(3.28)	(5.64)	(4.92)	(4.69)	(3.74)	(5.04)	(5.84)
Observations	6967	6967	6967	10914	10914	10914	26051	26051	26051
R^2	0.00	0.03	0.04	0.00	0.01	0.01	0.00	0.00	0.00
# Clusters	29	29	29	45	45	45	99	99	99

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A16: Dependent variable: RGGVY (= 100). Standard errors clustered by constituency. The sample is limited to 2002.

	Margin<1%			Margin<2%			Margin<5%		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BSP Win	0.01	-19.20	-19.09	-5.59	3.37	3.49	-1.81	0.92	0.83
	(7.02)	(15.03)	(14.73)	(5.75)	(10.26)	(10.37)	(3.69)	(6.92)	(6.90)
BSP Margin		20.77	18.69		-4.11	-3.36		-0.59	-0.36
		(13.08)	(13.79)		(4.59)	(6.57)		(1.24)	(2.04)
BSP Win * Margin			3.68			-1.60			-0.40
			(24.59)			(9.17)			(2.55)
Constant	30.06***	39.03***	38.13***	32.76***	28.34***	29.15***	31.60***	30.29***	30.80***
	(3.87)	(6.98)	(7.77)	(3.96)	(5.01)	(6.11)	(2.90)	(3.55)	(4.58)
Observations	7119	7119	7119	15879	15879	15879	36028	36028	36028
R^2	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.00	0.00
# Clusters	27	27	27	56	56	56	136	136	136

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A17: Dependent variable: RGGVY (= 100). Standard errors clustered by constituency. The sample is limited to 2007.

	Margin<1%			Margin<2%			Margin<5%		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BSP Win	4.39	29.87*	32.94**	0.69	13.82	13.76	0.23	1.46	1.38
	(7.74)	(15.03)	(14.98)	(8.02)	(10.81)	(10.39)	(5.75)	(8.77)	(8.58)
BSP Win * Share SC/ST	0.06	-0.03	-0.02	0.01	-0.03	-0.01	0.01	0.01	0.02
	(0.11)	(0.10)	(0.11)	(0.10)	(0.09)	(0.09)	(0.07)	(0.06)	(0.07)
Share SC/ST (%)	-0.19**	-0.13	-0.17*	-0.20***	-0.17**	-0.19***	-0.15***	-0.15***	-0.16***
	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)
BSP Margin		-27.07**	-9.68		-7.65	-2.82		-0.25	1.15
		(11.64)	(7.59)		(6.28)	(8.33)		(1.65)	(2.14)
BSP Win * Margin			-40.69			-10.28			-2.93
			(24.21)			(11.70)			(3.21)
Constant	33.17***	20.62***	28.68***	37.58***	30.57***	34.99***	34.02***	33.37***	37.02***
	(4.14)	(7.40)	(5.65)	(6.55)	(5.92)	(5.57)	(4.44)	(5.52)	(6.12)
Observations	6967	6967	6967	10914	10914	10914	26051	26051	26051
R^2	0.01	0.04	0.05	0.01	0.02	0.02	0.00	0.01	0.01
# Clusters	29	29	29	45	45	45	99	99	99

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A18: Dependent variable: RGGVY (= 100). The treatment (a BSP win) is interacted with the share of SC in the village. Standard errors clustered by constituency. The sample is limited to 2002.

	Margin<1%			Margin<2%			Margin<5%		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BSP Win	0.44	-17.61	-17.57	-7.87	-0.39	-0.36	-0.63	1.43	1.23
	(7.88)	(15.21)	(15.10)	(6.85)	(10.05)	(10.09)	(4.52)	(7.31)	(7.27)
BSP Win * Share SC/ST	-0.08	-0.07	-0.07	0.09	0.10	0.10	-0.03	-0.03	-0.03
	(0.09)	(0.10)	(0.10)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)
Share SC/ST (%)	-0.30***	-0.30***	-0.30***	-0.33***	-0.33***	-0.33***	-0.23***	-0.23***	-0.23***
	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
BSP Margin		19.32	17.67		-3.50	-2.99		-0.43	-0.08
		(12.93)	(14.36)		(4.53)	(6.47)		(1.22)	(2.03)
BSP Win * Margin			2.93			-1.09			-0.62
			(24.52)			(9.05)			(2.53)
Constant	37.66***	45.79***	45.10***	40.80***	36.93***	37.50***	36.98***	35.96***	36.80***
	(5.06)	(7.27)	(8.04)	(4.97)	(5.47)	(6.22)	(3.57)	(4.05)	(4.94)
Observations	7119	7119	7119	15879	15879	15879	36028	36028	36028
R^2	0.02	0.03	0.03	0.02	0.02	0.02	0.01	0.01	0.01
# Clusters	27	27	27	56	56	56	136	136	136

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A19: Dependent variable: RGGVY (= 100). The treatment (a BSP win) is interacted with the share of SC in the village. Standard errors clustered by constituency. The sample is limited to 2007.

	Margin<3%				Margin<4%			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BSP Win	-0.50	-0.46	3.00	2.96	-1.28	-1.28	3.09	2.96
	(3.87)	(3.85)	(6.11)	(6.11)	(3.45)	(3.45)	(5.70)	(5.68)
2007 Election		-0.95	-0.86	-0.58		0.19	0.16	0.26
		(3.33)	(3.30)	(3.27)		(2.78)	(2.77)	(2.74)
BSP Margin			-1.18	-0.25			-1.16	-0.61
			(1.88)	(2.74)			(1.31)	(1.99)
BSP Win * Margin				-1.83				-1.01
				(3.75)				(2.59)
Constant	31.26***	31.82***	30.05***	31.23***	31.64***	31.52***	29.43***	30.38***
	(2.93)	(3.85)	(3.90)	(4.23)	(2.70)	(3.40)	(3.69)	(4.22)
Observations	40121	40121	40121	40121	50797	50797	50797	50797
R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
# Clusters	142	142	142	142	174	174	174	174

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A20: Dependent variable: RGGVY (= 100). The treatment (a BSP win) is interacted with the share of SC in the village. Standard errors clustered by constituency. Different set of bandwidths.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SC Win	1.05	1.05	-1.94	-2.22	2.09	2.09	-0.87	-1.27
	(2.63)	(2.63)	(3.00)	(3.11)	(3.22)	(3.22)	(3.60)	(3.78)
2007 Election		-0.17	-0.31	-0.31		-0.14	-0.28	-0.27
		(0.57)	(0.61)	(0.62)		(0.57)	(0.61)	(0.62)
SC Margin			0.21*	0.20			0.21*	0.19
			(0.12)	(0.13)			(0.12)	(0.13)
SC Win * Margin				0.07				0.10
				(0.27)				(0.27)
SC Win * Share SC					0.00	0.00	-0.00	-0.00
					(0.05)	(0.05)	(0.05)	(0.05)
Share SC (%)					-0.19***	-0.19***	-0.19***	-0.19***
					(0.02)	(0.02)	(0.02)	(0.02)
Constant	30.48***	30.57***	32.14***	32.04***	34.78***	34.85***	36.45***	36.30***
	(1.29)	(1.36)	(1.69)	(1.77)	(1.41)	(1.48)	(1.80)	(1.88)
Observations	181791	181791	181791	181791	181791	181791	181791	181791
R^2	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
# Clusters	399	399	399	399	399	399	399	399

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A21: RDD analysis with SC victory effects (instead of BSP): 1%, 2%, 5% bandwidths. Dependent variable: RGGVY (if present, RGGVY= 100). The standard errors are clustered by constituency.

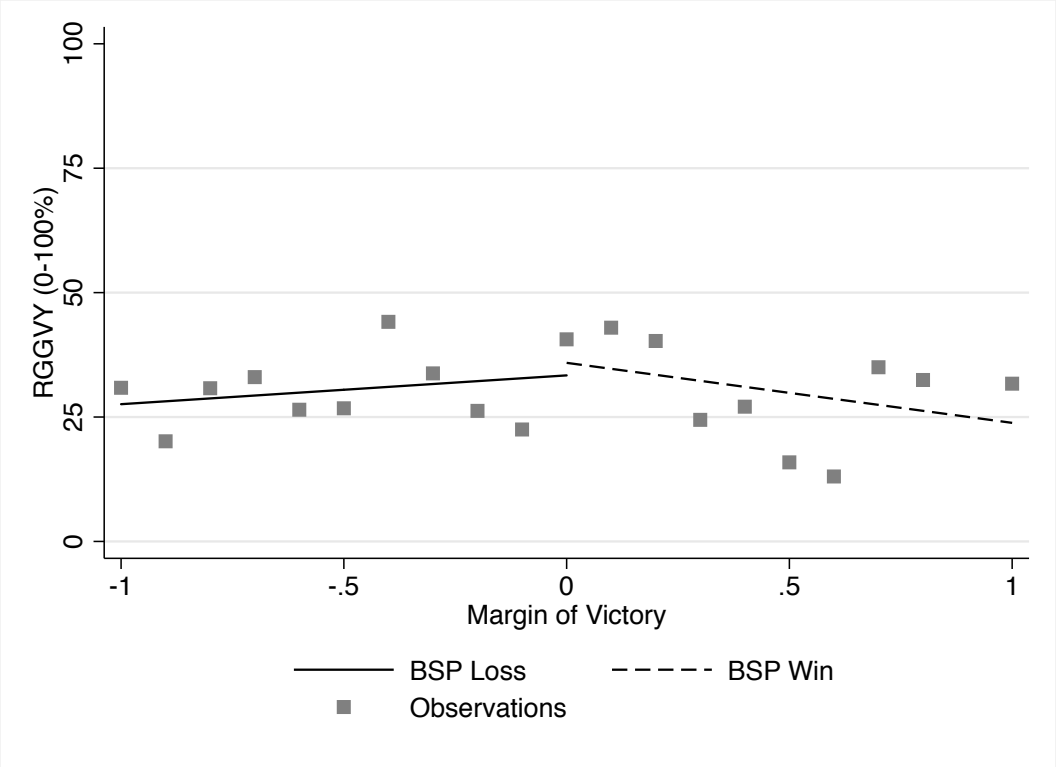

Figure A7: Regression discontinuity graph with a 1% margin. Observations are binned by slides of 0.1 (i.e. from -1 to -0.9, from -0.9 to -0.8, ..., from 0.9 to 1). Within each bin, we take the share of villages that have benefited from RGGVY. These are the observations plotted on the x- and y-axis, respectively. We then fit two linear regressions on either side of the cutoff.

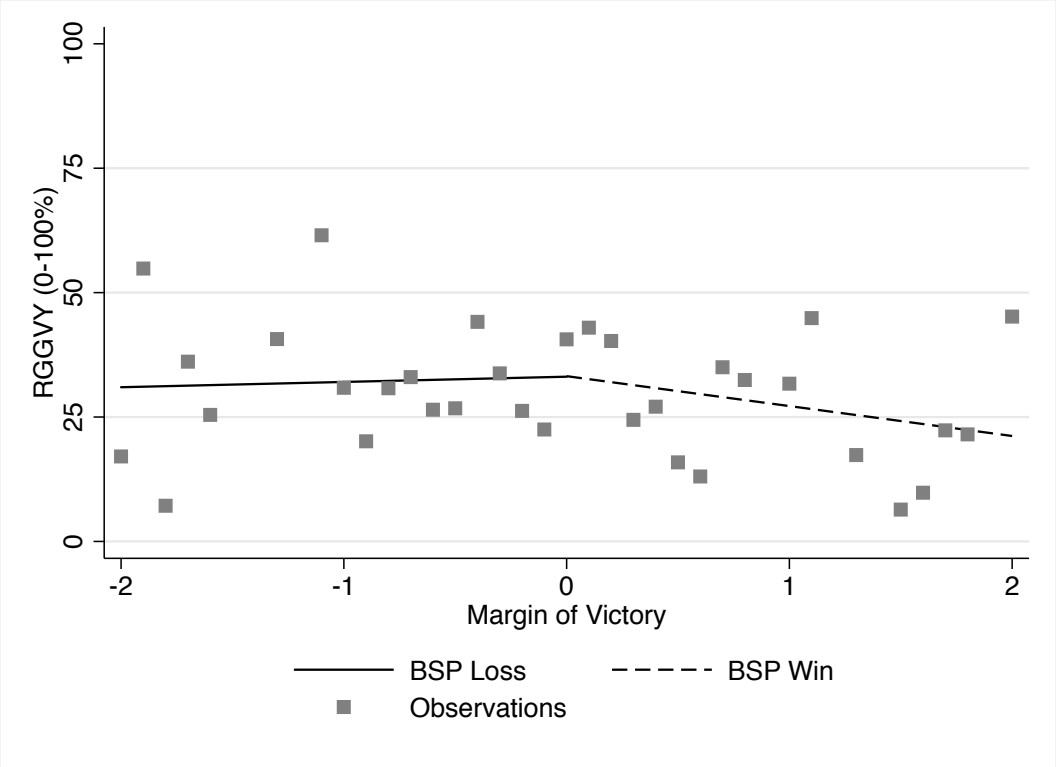

Figure A8: Regression discontinuity graph with a 2% margin. Observations are binned by slides of 0.1 (i.e. from -2 to -1.9, from -1.9 to -1.8, ..., from 1.9 to 2). Within each bin, we take the share of villages that have benefited from RGGVY. These are the observations plotted on the x- and y-axis, respectively. We then fit two linear regressions on either side of the cutoff.

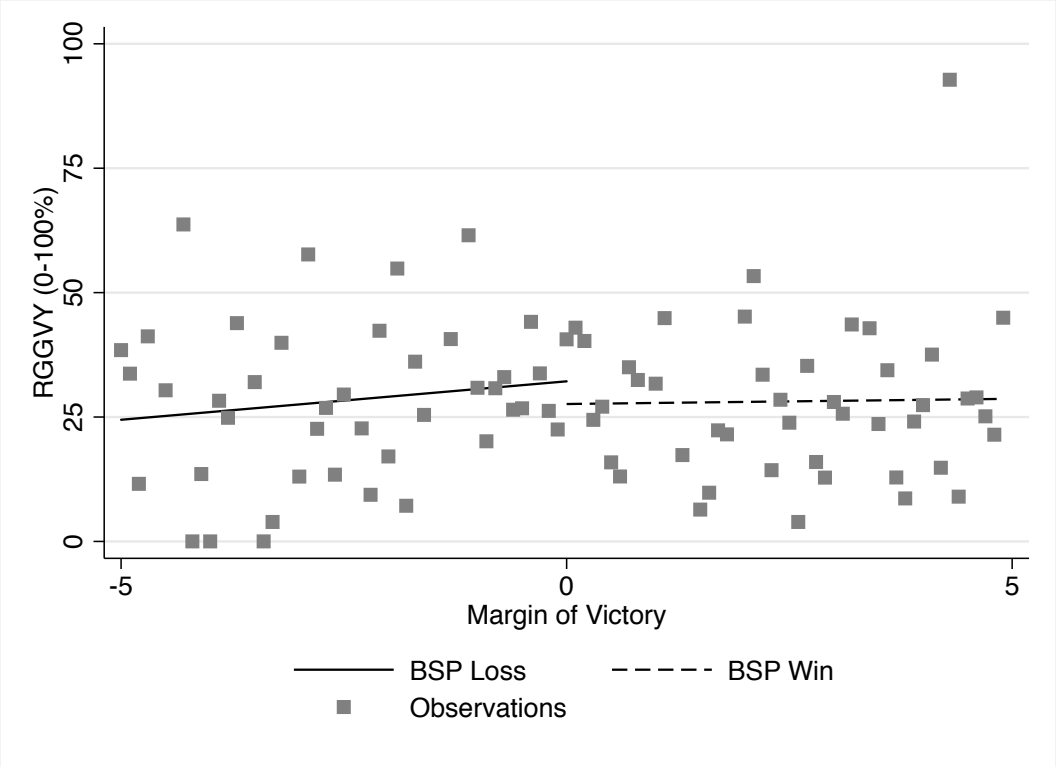

Figure A9: Regression discontinuity graph with a 5% margin. Observations are binned by slides of 0.1 (i.e. from -5 to -4.9, from -4.9 to -4.8, ..., from 4.9 to 5). Within each bin, we take the share of villages that have benefited from RGGVY. These are the observations plotted on the x- and y-axis, respectively. We then fit two linear regressions on either side of the cutoff.

A7 Regression Discontinuity: Conditioning on Reservation Status

- Table A22 conditions the effect of a BSP win on the constituency reservation status (SC versus general). The samples are based on the regression discontinuity thresholds.
- Table A23 conditions the effect of a BSP win on the constituency reservation status (SC versus general). All available observations are used to produce the estimates.
- Table A24 splits the analysis between reserved and non-reserved constituencies. As a result, the effect of a BSP win is interacted with the SC share in the village population, the margin of victory, and the reservation status.

	Margin<1%				Margin<2%				Margin<5%			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	
BSP Win	1.81	1.81	2.98	3.64	-6.72	-6.95	3.12	2.83	-1.77	-1.77	0.51	
	(5.86)	(5.86)	(12.40)	(13.49)	(5.53)	(5.59)	(8.15)	(8.10)	(4.04)	(4.04)	(5.86)	(
BSP Win * Share SC	-0.01	-0.01	-0.02	-0.01	0.03	0.03	0.02	0.03	-0.05	-0.05	-0.05	-
	(0.09)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(
Share SC (%)	-0.25***	-0.25***	-0.25***	-0.26***	-0.25***	-0.26***	-0.25***	-0.25***	-0.18***	-0.18***	-0.18***	-0
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(
Reserved Constituency	2.63	2.63	2.48	2.93	-6.58	-7.93	-5.91	-6.76	-4.85	-5.02	-4.93	-
	(4.42)	(4.40)	(4.51)	(4.10)	(5.29)	(5.61)	(5.97)	(5.79)	(5.58)	(5.82)	(5.94)	(
Reserved*BSP Win	17.97	17.97	17.72	15.98	16.48	17.70	16.32	17.59	11.74	11.87	12.06	1
	(11.26)	(11.53)	(12.38)	(14.31)	(13.01)	(12.78)	(12.34)	(11.99)	(7.82)	(7.98)	(7.93)	(
2007 Election		0.00	-0.04	0.03		-4.31	-3.85	-3.02		-0.46	-0.36	-
		(4.82)	(4.98)	(4.95)		(4.60)	(4.52)	(4.57)		(2.56)	(2.57)	(
BSP Margin			-1.28	2.69			-4.91	-1.99			-0.48	
			(10.92)	(7.25)			(3.67)	(5.05)			(0.96)	(
BSP Win * Margin				-8.43				-6.02				-
				(22.20)				(7.23)				(
Constant	34.71***	34.71***	34.22***	35.79***	40.67***	43.63***	37.87***	40.54***	36.41***	36.71***	35.46***	38
	(4.26)	(4.91)	(5.91)	(5.97)	(4.49)	(5.85)	(6.09)	(6.42)	(3.21)	(3.78)	(4.39)	(
Observations	14086	14086	14086	14086	26793	26793	26793	26793	62079	62079	62079	6
R^2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.01	0.01	0.01	
# Clusters	55	55	55	55	97	97	97	97	200	200	200	

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A22: Dependent variable: RGGVY (= 100). Standard errors clustered by constituency. Additional controls based on reservation status. Regression discontinuity framework.

	(1)	(2)	(3)	(4)
BSP Win	-1.21	-1.46	-0.15	-0.15
	(2.16)	(2.25)	(3.63)	(3.62)
BSP Win * Share SC	-0.03	-0.03	-0.05	-0.05
	(0.03)	(0.03)	(0.03)	(0.04)
Share SC (%)	-0.18***	-0.18***	-0.15***	-0.15***
	(0.03)	(0.03)	(0.03)	(0.03)
Reserved Constituency	4.65	4.74	-2.50	-2.60
	(4.15)	(4.16)	(4.27)	(4.32)
Reserved*BSP Win	-4.91	-5.07	2.30	2.50
	(4.99)	(5.03)	(4.68)	(4.84)
Triple Interaction	0.06	0.06	0.06	0.06
	(0.05)	(0.05)	(0.05)	(0.05)
2007 Election		0.90	0.30	0.30
		(0.60)	(1.19)	(1.18)
BSP Margin			-0.06	-0.00
			(0.19)	(0.26)
BSP Win * Margin				-0.12
				(0.33)
Constant	35.46***	35.09***	34.54***	34.92***
	(1.73)	(1.67)	(2.75)	(3.07)
Observations	193114	193114	131340	131340
R^2	0.01	0.01	0.01	0.01
# Clusters	402	402	340	340

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A23: Dependent variable: RGGVY (= 100). Standard errors clustered by constituency. Additional controls based on reservation status. The sample includes all available observations.

	Non-Reserved				Reserved			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BSP Win	-1.10	-1.22	-1.42	-0.99	-6.51	-7.65	6.89	10.99
	(2.16)	(2.29)	(3.77)	(3.79)	(4.32)	(5.17)	(7.97)	(8.22)
BSP Win * Share SC	-0.03	-0.03	-0.04	-0.04	0.04	0.04	-0.01	-0.02
	(0.03)	(0.03)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Share SC (%)	-0.17***	-0.17***	-0.16***	-0.16***	-0.19***	-0.19***	-0.13*	-0.13*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)	(0.06)	(0.07)	(0.07)
2007 Election		0.41	0.29	0.31		2.71	-0.11	-0.61
		(0.48)	(1.30)	(1.28)		(2.24)	(2.84)	(2.82)
BSP Margin			0.03	0.18			-0.37	-1.77*
			(0.21)	(0.23)			(0.45)	(1.03)
BSP Win * Margin				-0.39				1.83*
				(0.33)				(1.10)
Constant	35.36***	35.19***	35.31***	36.35***	40.49***	39.62***	29.98***	22.97***
	(1.74)	(1.64)	(2.90)	(2.96)	(3.97)	(3.61)	(4.50)	(6.36)
Observations	147718	147718	98356	98356	45396	45396	32984	32984
R^2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
# Clusters	314	314	259	259	89	89	82	82

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A24: Dependent variable: RGGVY (= 100). Standard errors clustered by constituency. Sample split by reservation status.

A8 Regional Samples

- Tables A25-A28 show the main estimation results by region (West, Central, East, and Bundelkhand).

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.31***	-0.31***	-0.28***	-0.29***	-0.30***	-0.26***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		2.77***				2.37***
		(0.48)				(0.44)
Domestic Electricity (2001)			-0.14***			-0.14***
			(0.01)			(0.01)
Population (log)				-7.02***		-6.70***
				(0.73)		(0.75)
Pucca Road					-3.79***	-1.50**
					(0.75)	(0.74)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	27022	25116	27022	27022	26970	25091
R^2	0.02	0.03	0.05	0.05	0.02	0.08
# Clusters	147	146	147	147	147	146

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A25: Dependent variable: RGGVY (= 100). The sample is limited to districts in the Western Region. Standard errors clustered by constituency.

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.14***	-0.13***	-0.10***	-0.17***	-0.13***	-0.13***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Distance (log)		1.15**				0.77*
		(0.49)				(0.45)
Domestic Electricity (2001)			-0.22***			-0.19***
			(0.02)			(0.02)
Population (log)				-11.55***		-10.05***
				(1.31)		(1.23)
Pucca Road					-7.98***	-2.49**
					(1.33)	(1.07)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	15098	13930	15098	15098	15017	13868
R^2	0.00	0.00	0.06	0.06	0.01	0.11
# Clusters	77	77	77	77	77	77

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A26: Dependent variable: RGGVY (= 100). The sample is limited to districts in the Central Region. Standard errors clustered by constituency.

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.17***	-0.17***	-0.17***	-0.17***	-0.17***	-0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		0.90***				0.73***
		(0.27)				(0.27)
Domestic Electricity (2001)			-0.12***			-0.11***
			(0.01)			(0.01)
Population (log)				-5.15***		-4.52***
				(0.50)		(0.47)
Pucca Road					-5.03***	-2.32***
					(0.67)	(0.59)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	49976	47511	49976	49976	49817	47422
R^2	0.01	0.01	0.02	0.02	0.01	0.04
# Clusters	160	160	160	160	160	160

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A27: Dependent variable: RGGVY (= 100). The sample is limited to districts in the Eastern Region. Standard errors clustered by constituency.

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.33***	-0.33***	-0.26***	-0.27***	-0.33***	-0.23***
	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)
Distance (log)		2.00*				2.12*
		(1.14)				(1.14)
Domestic Electricity (2001)			-0.23***			-0.20***
			(0.02)			(0.02)
Population (log)				-9.85***		-9.40***
				(1.17)		(1.39)
Pucca Road					-8.85***	-3.66*
					(2.01)	(1.93)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	4461	4126	4461	4461	4392	4073
R^2	0.02	0.02	0.08	0.08	0.03	0.14
# Clusters	24	24	24	24	24	24

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A28: Dependent variable: RGGVY (= 100). The sample is limited to districts in the Bundelkhand Region (south). Standard errors clustered by constituency.

A9 Additional Results

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	0.16*	0.14	0.13	0.12	0.11	0.04
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)
Distance (log)		-3.80*				-3.26
		(2.20)				(2.19)
Domestic Electricity (2001)			0.31***			0.23***
			(0.04)			(0.03)
Population (log)				40.66***		35.26***
				(2.01)		(2.07)
Pucca Road					79.13***	62.78***
					(7.29)	(7.79)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	95963	90236	95963	95963	95791	90110
R^2	0.00	0.00	0.00	0.02	0.01	0.03
# Clusters	402	401	402	402	402	401

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A29: Dependent variable: average hours of power supply per day (rescaled to $0 - 2400$ for readability). Standard errors clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.19***	-0.18***	-0.17***	-0.18***	-0.18***	-0.16***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Wealth Index	-3.78***	-3.93***	-3.94***	-3.36***	-3.80***	-3.67***
	(0.69)	(0.69)	(0.68)	(0.69)	(0.70)	(0.69)
Literacy Rate (%)	-0.45***	-0.45***	-0.43***	-0.50***	-0.45***	-0.48***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
# Coop Commercial Banks	-12.87***	-12.54***	-12.94***	-5.41***	-12.20***	-5.79***
	(1.41)	(1.45)	(1.39)	(1.21)	(1.37)	(1.25)
Irrigated Land (log)	-1.14***	-1.17***	-1.06***	-0.31	-1.08***	-0.36*
	(0.21)	(0.21)	(0.21)	(0.22)	(0.21)	(0.21)
Mean Light	-1.71***	-1.63***	-1.72***	-1.58***	-1.68***	-1.51***
	(0.15)	(0.15)	(0.15)	(0.14)	(0.15)	(0.15)
Distance (log)		2.02***				2.16***
		(0.47)				(0.44)
Domestic Electricity (2001)			-0.15***			-0.14***
			(0.01)			(0.01)
Population (log)				-6.08***		-5.30***
				(0.48)		(0.46)
Pucca Road					-4.70***	-2.19***
					(0.88)	(0.83)
Constant	69.88***	64.87***	73.68***	111.15***	72.22***	104.98***
	(2.91)	(3.17)	(2.82)	(4.21)	(2.88)	(4.46)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.05	0.05	0.07	0.06	0.05	0.09
# Clusters	402	401	402	402	402	401

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.22***	-0.22***	-0.21***	-0.22***	-0.22***	-0.20***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Wealth Index	-1.85***	-1.86***	-1.80***	-1.61***	-1.87***	-1.59***
	(0.24)	(0.24)	(0.23)	(0.23)	(0.23)	(0.22)
Literacy Rate (%)	-0.41***	-0.42***	-0.37***	-0.44***	-0.41***	-0.41***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
# Coop Commercial Banks	-12.60***	-12.28***	-12.97***	-5.54***	-12.10***	-6.08***
	(1.04)	(1.07)	(1.06)	(0.95)	(1.03)	(0.99)
Irrigated Land (log)	-1.17***	-1.18***	-1.04***	-0.24**	-1.11***	-0.20**
	(0.11)	(0.12)	(0.11)	(0.10)	(0.11)	(0.10)
Mean Light	-0.98***	-0.92***	-0.94***	-0.89***	-0.98***	-0.81***
	(0.10)	(0.10)	(0.09)	(0.10)	(0.09)	(0.10)
Distance (log)		0.86***				0.69***
		(0.20)				(0.20)
Domestic Electricity (2001)			-0.14***			-0.13***
			(0.01)			(0.01)
Population (log)				-6.47***		-5.81***
				(0.39)		(0.38)
Pucca Road					-4.80***	-2.09***
					(0.45)	(0.41)
Constant	66.54***	64.79***	68.28***	109.13***	69.42***	106.33***
	(1.73)	(1.82)	(1.66)	(3.47)	(1.79)	(3.30)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.03	0.03	0.05	0.05	0.03	0.07
# Clusters	402	401	402	402	402	401
Constituency FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A30: Dependent variable: RGGVY (if present, RGGVY= 100). The standard errors are clustered by constituency. Additional control variables (pre-RGGVY) for wealth and wealth-related confounders: literacy rate (%), number of cooperative commercial banks, irrigated land area (logarithmized), population (logarithmized), average decadal (1995-2004) nighttime luminosity in digital number on a 0-64 scale, with higher values indicating more llight. All variables are from the 2001 Census of India, except night lights are from NOAA satellites.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.18***	-0.17***	-0.16***	-0.18***	-0.17***	-0.15***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Margin of Victory	-0.41	-0.41	-0.38	-0.46	-0.37	-0.42
	(0.38)	(0.39)	(0.37)	(0.37)	(0.37)	(0.36)
Margin*Share SC	0.00	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance (log)		2.52***				2.53***
		(0.73)				(0.70)
Domestic Electricity (2001)			-0.16***			-0.14***
			(0.02)			(0.01)
Population (log)				-6.73***		-5.78***
				(0.84)		(0.84)
Pucca Road					-6.78***	-3.00*
					(1.59)	(1.62)
Constant	36.31***	30.49***	40.99***	82.43***	40.45***	76.12***
	(3.66)	(4.04)	(3.80)	(7.69)	(3.50)	(7.58)
Observations	25556	23877	25556	25556	25455	23829
R^2	0.01	0.01	0.03	0.03	0.01	0.06
# Clusters	98	98	98	98	98	98

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.20***	-0.20***	-0.19***	-0.20***	-0.19***	-0.19***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Margin*Share SC	-0.00	-0.00	0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance (log)		1.28***				1.08**
		(0.44)				(0.42)
Domestic Electricity (2001)			-0.14***			-0.13***
			(0.02)			(0.01)
Population (log)				-7.05***		-6.12***
				(0.73)		(0.66)
Pucca Road					-6.35***	-3.17***
					(0.89)	(0.77)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	25556	23877	25556	25556	25455	23829
R^2	0.01	0.01	0.03	0.04	0.01	0.06
# Clusters	98	98	98	98	98	98

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A31: Safe vs. marginal seats. Dependent variable: RGGVY (if present, RGGVY= 100). The standard errors are clustered by constituency. The sample is limited to cases in which a BSP member won.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.16***	-0.16***	-0.14***	-0.16***	-0.15***	-0.13***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		3.15***				3.26***
		(0.48)				(0.46)
Domestic Electricity (2001)			-0.16***			-0.14***
			(0.01)			(0.01)
Population (log)				-6.52***		-5.66***
				(0.50)		(0.48)
Pucca Road					-6.48***	-3.61***
					(1.00)	(0.96)
Constant	34.95***	27.38***	40.14***	79.87***	39.10***	73.07***
	(1.31)	(1.58)	(1.45)	(3.95)	(1.33)	(4.09)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.03	0.03	0.01	0.06
# Clusters	402	401	402	402	402	401

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.21***	-0.21***	-0.19***	-0.20***	-0.21***	-0.19***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance (log)		1.34***				1.10***
		(0.22)				(0.21)
Domestic Electricity (2001)			-0.15***			-0.13***
			(0.01)			(0.01)
Population (log)				-6.64***		-5.94***
				(0.40)		(0.38)
Pucca Road					-5.41***	-2.31***
					(0.47)	(0.42)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.04	0.04	0.01	0.06
# Clusters	402	401	402	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A32: Dependent variable: RGGVY (if present, RGGVY= 100). We report the effect of scheduled castes (instead of the combined number of scheduled tribes and scheduled castes). The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share ST (%)	-0.31***	-0.31***	-0.32***	-0.33***	-0.32***	-0.35***
	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)
Distance (log)		3.14***				3.26***
		(0.47)				(0.45)
Domestic Electricity (2001)			-0.16***			-0.15***
			(0.01)			(0.01)
Population (log)				-6.59***		-5.68***
				(0.50)		(0.48)
Pucca Road					-6.88***	-3.92***
					(1.00)	(0.97)
Constant	31.30***	23.88***	37.06***	76.86***	35.90***	70.62***
	(1.21)	(1.46)	(1.36)	(3.80)	(1.25)	(3.96)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.00	0.01	0.03	0.03	0.01	0.06
# Clusters	402	401	402	402	402	401

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share ST (%)	-0.15**	-0.14*	-0.17**	-0.15**	-0.16**	-0.17**
	(0.07)	(0.07)	(0.07)	(0.06)	(0.07)	(0.07)
Distance (log)		1.40***				1.15***
		(0.22)				(0.21)
Domestic Electricity (2001)			-0.15***			-0.14***
			(0.01)			(0.01)
Population (log)				-6.72***		-5.97***
				(0.40)		(0.38)
Pucca Road					-5.76***	-2.58***
					(0.48)	(0.42)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	96557	90683	96557	96557	96196	90454
R^2	0.00	0.00	0.03	0.03	0.00	0.05
# Clusters	402	401	402	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A33: Dependent variable: RGGVY (if present, RGGVY= 100). We report the effect of scheduled tribes (instead of the combined number of scheduled tribes and scheduled castes). The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC in 2001 (%)	-0.18***	-0.17***	-0.16***	-0.17***	-0.17***	-0.15***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		3.15***				3.25***
		(0.47)				(0.46)
Domestic Electricity (2001)			-0.16***			-0.14***
			(0.01)			(0.01)
Population (log)				-6.55***		-5.69***
				(0.50)		(0.49)
Pucca Road					-6.47***	-3.56***
					(0.99)	(0.96)
Constant	35.40***	27.84***	40.59***	80.62***	39.56***	73.83***
	(1.33)	(1.60)	(1.47)	(3.97)	(1.36)	(4.12)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.03	0.03	0.01	0.06
# Clusters	402	401	402	402	402	401

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC in 2001 (%)	-0.21***	-0.21***	-0.20***	-0.21***	-0.21***	-0.19***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance (log)		1.34***				1.10***
		(0.22)				(0.21)
Domestic Electricity (2001)			-0.15***			-0.13***
			(0.01)			(0.01)
Population (log)				-6.67***		-5.96***
				(0.40)		(0.38)
Pucca Road					-5.44***	-2.32***
					(0.47)	(0.42)
Constituency FE	✓	✓	✓	✓	✓	✓
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.04	0.04	0.01	0.06
# Clusters	402	401	402	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A34: Dependent variable: RGGVY (if present, RGGVY= 100). The data for the share of SC come from the 2001 Census instead of the 2011 one. The correlation between 2001 and 2011 share of SC is 0.92. The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.30***	-0.30***	-0.24***	0.05	-0.26***	0.08
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Share SC (square)	0.00**	0.00***	0.00	-0.00***	0.00*	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance (log)		3.19***				3.20***
		(0.47)				(0.45)
Domestic Electricity (2001)			-0.16***			-0.14***
			(0.01)			(0.01)
Population (log)				-7.20***		-6.36***
				(0.49)		(0.48)
Pucca Road					-6.33***	-3.61***
					(0.98)	(0.96)
Constant	36.73***	29.12***	41.43***	82.65***	40.46***	76.11***
	(1.37)	(1.74)	(1.50)	(3.93)	(1.42)	(4.06)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.03	0.03	0.01	0.06
# Clusters	402	401	402	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.46***	-0.46***	-0.42***	-0.13***	-0.44***	-0.12***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
Share SC (square)	0.00***	0.00***	0.00***	-0.00***	0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance (log)		1.36***				1.10***
		(0.22)				(0.21)
Domestic Electricity (2001)			-0.14***			-0.13***
			(0.01)			(0.01)
Population (log)				-6.87***		-6.15***
				(0.40)		(0.39)
Pucca Road					-5.09***	-2.34***
					(0.46)	(0.42)
Constant	38.98***	35.84***	43.44***	82.71***	42.11***	81.16***
	(0.43)	(0.63)	(0.53)	(2.88)	(0.58)	(2.88)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.04	0.04	0.01	0.06
# Clusters	402	401	402	402	402	401
Constituency FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A35: Quadratic effect of SC share. Dependent variable: RGGVY (if present, RGGVY= 100). The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Distance (log)		0.15***				0.16***
		(0.02)				(0.02)
Domestic Electricity (2001)			-0.01***			-0.01***
			(0.00)			(0.00)
Population (log)				-0.30***		-0.26***
				(0.02)		(0.02)
Pucca Road					-0.30***	-0.17***
					(0.05)	(0.05)
Constant	-0.59***	-0.96***	-0.35***	1.43***	-0.40***	1.12***
	(0.06)	(0.08)	(0.06)	(0.18)	(0.06)	(0.19)
Observations	96557	90683	96557	96557	96196	90454
# Clusters	402	401	402	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A36: Logit specification. Dependent variable: RGGVY (if present, RGGVY= 1). Note: fixed effect versions are computationally too intensive, given the large number of parameters. The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC [0,1]	-0.28	-0.26	-0.28	-0.30	-0.33	-0.31
	(0.21)	(0.20)	(0.21)	(0.21)	(0.21)	(0.20)
Distance (log)		-0.53***				-0.55***
		(0.06)				(0.06)
Domestic Electricity (2001)			0.00			-0.00
			(0.00)			(0.00)
Population (log)				0.30***		0.26***
				(0.05)		(0.04)
Pucca Road					0.51***	0.41***
					(0.14)	(0.13)
Constant	2.58***	3.83***	2.57***	0.52	2.26***	1.80***
	(0.14)	(0.25)	(0.14)	(0.33)	(0.16)	(0.40)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.00	0.02	0.00	0.01	0.00	0.04
# Clusters	402	401	402	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC [0,1]	0.02	-0.02	0.01	0.01	0.01	-0.03
	(0.12)	(0.13)	(0.13)	(0.12)	(0.13)	(0.13)
Distance (log)		-0.44***				-0.43***
		(0.05)				(0.05)
Domestic Electricity (2001)			0.00***			0.00**
			(0.00)			(0.00)
Population (log)				0.08***		0.08***
				(0.02)		(0.02)
Pucca Road					-0.01	-0.03
					(0.04)	(0.04)
Constant	2.51***	3.54***	2.48***	1.94***	2.52***	2.98***
	(0.03)	(0.11)	(0.03)	(0.11)	(0.04)	(0.14)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.00	0.03	0.00	0.00	0.00	0.03
# Clusters	402	401	402	402	402	401
Constituency FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A37: Effect of SC share on night-time lighting. Dependent variable: average night-time light. Note: the share of SC is rescaled to the $[0, 1]$ interval to make point estimates more readable. The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)
Share SC (%)	0.11***	0.10***	0.10***	0.10***	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		2.46***			2.29***
		(0.50)			(0.47)
Domestic Electricity (2001)			0.03***		0.01
			(0.01)		(0.01)
Population (log)				9.72***	9.60***
				(0.48)	(0.50)
Constant	63.78***	57.30***	62.86***	-3.22	-8.59*
	(1.76)	(2.51)	(1.79)	(4.37)	(4.88)
Observations	96196	90454	96196	96196	90454
R^2	0.00	0.01	0.00	0.05	0.05
# Clusters	402	401	402	402	401

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)
Share SC (%)	0.06***	0.06***	0.06***	0.06***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance (log)		0.46**			0.62***
		(0.23)			(0.22)
Domestic Electricity (2001)			0.03***		0.01***
			(0.00)		(0.00)
Population (log)				7.59***	7.48***
				(0.25)	(0.25)
Constant	64.77***	63.09***	63.92***	12.49***	10.96***
	(0.23)	(0.58)	(0.27)	(1.74)	(1.89)
Observations	96196	90454	96196	96196	90454
R^2	0.00	0.00	0.00	0.04	0.04
# Clusters	402	401	402	402	401
Constituency FE	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A38: Dependent variable: pucca road (if present, pucca= 100). The standard errors are clustered by constituency.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.18***	-0.17***	-0.16***	-0.17***	-0.17***	-0.15***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Distance (log)		3.15***				3.26***
		(0.82)				(0.77)
Domestic Electricity (2001)			-0.16***			-0.14***
			(0.02)			(0.02)
Population (log)				-6.53***		-5.67***
				(0.93)		(0.92)
Pucca Road					-6.48***	-3.59*
					(2.00)	(1.92)
Constant	35.43***	27.86***	40.64***	80.42***	39.60***	73.64***
	(2.70)	(2.72)	(2.95)	(7.36)	(2.54)	(7.91)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.01	0.03	0.03	0.01	0.06
# Clusters	70	70	70	70	70	70

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.21***	-0.21***	-0.20***	-0.20***	-0.21***	-0.19***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		1.35***				1.10***
		(0.26)				(0.24)
Domestic Electricity (2001)			-0.15***			-0.13***
			(0.02)			(0.01)
Population (log)				-6.64***		-5.93***
				(0.78)		(0.73)
Pucca Road					-5.44***	-2.34***
					(0.72)	(0.55)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.21	0.21	0.23	0.23	0.21	0.25
# Clusters	70	70	70	70	70	70
Constituency FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A39: Dependent variable: RGGVY (if present, RGGVY= 100). The standard errors are clustered by district.

	Unelectrified in 2001					Electrified in 2001				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share SC (%)	-0.20***	-0.20***	-0.19***	-0.20***	-0.19***	-0.17***	-0.17***	-0.17***	-0.17***	-0.17***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		1.49***			1.34***		0.54***			0.52***
		(0.31)			(0.30)		(0.19)			(0.18)
Population (log)			-6.08***		-5.86***			-4.73***		-4.65***
			(0.76)		(0.78)			(0.69)		(0.68)
Pucca Road				-6.11***	-3.18***				-1.67**	-0.13
				(0.73)	(0.61)				(0.69)	(0.62)
Observations	61950	58245	61950	61724	58104	34605	32435	34605	34471	32348
R^2	0.23	0.23	0.25	0.23	0.25	0.23	0.23	0.24	0.23	0.25
# Clusters	70	70	70	70	70	70	70	70	70	70
Constituency FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A40: Dependent variable: RGGVY (if present, RGGVY= 100). All models estimated with constituency fixed effects. The standard errors are clustered by district.

Pooled

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.09***	-0.09***	-0.09***	-0.09***	-0.09***	-0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Distance (log)		-2.13***				-2.15***
		(0.34)				(0.33)
Domestic Electricity (2001)			0.00			0.00
			(0.01)			(0.01)
Population (log)				-0.73		-0.96*
				(0.47)		(0.49)
Pucca Road					0.86	1.44
					(1.42)	(1.51)
Constant	25.57***	30.66***	25.53***	30.59***	25.05***	36.30***
	(1.46)	(1.73)	(1.53)	(2.83)	(1.66)	(3.13)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.01	0.02	0.01	0.01	0.01	0.02
# Clusters	70	70	70	70	70	70

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Constituency fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Share SC (%)	-0.08***	-0.09***	-0.08***	-0.08***	-0.08***	-0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Distance (log)		-1.52***				-1.54***
		(0.14)				(0.14)
Domestic Electricity (2001)			0.01**			0.01**
			(0.00)			(0.00)
Population (log)				-0.77***		-0.97***
				(0.24)		(0.25)
Pucca Road					0.74***	1.16***
					(0.22)	(0.22)
Observations	96557	90683	96557	96557	96196	90454
R^2	0.30	0.29	0.30	0.30	0.30	0.30
# Clusters	70	70	70	70	70	70
Constituency FE	✓	✓	✓	✓	✓	✓

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A41: Dependent variable: household electrification in 2011 (0-100 percent). The standard errors are clustered by district.

Spatial Autoregressive Models (Part I)

District	Point estimate (Share SC)	Standard error	t-value
Agra	-0.37	0.06	-6.44
Aligarh	-0.41	0.06	-6.29
Allahabad	-0.27	0.05	-5.69
Ambedkar Nagar	-0.07	0.03	-2.21
Auraiya	-0.32	0.08	-4.05
Azamgarh	-0.24	0.03	-7.32
Baghpat	-	-	-
Bahraich	-0.15	0.10	-1.50
Ballia	-0.22	0.05	-4.23
Balrampur	-0.23	0.08	-3.00
Banda	-0.50	0.12	-4.31
Barabanki	-0.17	0.06	-2.83
Bareilly	-0.37	0.05	-7.50
Basti	-0.16	0.05	-3.38
Bijnor	-0.21	0.03	-7.63
Budaun	-0.47	0.04	-10.87
Bulandshahar	-0.25	0.06	-4.32
Chandauli	-0.01	0.05	-0.12
Chitrakoot	-0.36	0.09	-3.97
Deoria	-0.15	0.04	-3.56
Etah	-0.81	0.06	-13.75
Etawah	-0.42	0.08	-4.93
Faizabad	-0.33	0.08	-4.15
Farrukhabad	-0.42	0.09	-4.82
Fatehpur	-0.18	0.08	-2.15
Firozabad	-0.63	0.07	-8.73
Gautam Buddha Nagar	-0.28	0.12	-2.29
Ghaziabad	0.07	0.05	1.45
Ghazipur	-0.08	0.02	-3.92
Gonda	-0.55	0.08	-6.49
Gorakhpur	-0.17	0.03	-5.45
Hamirpur	-0.61	0.16	-3.69
Hardoi	-0.23	0.05	-4.43
Hathras	-0.29	0.07	-4.44

Table A42: Estimates from a spatial autoregressive model (part I). Estimates computed state-by-state due to the size of the spatial correlation matrix. Models could not converge in a few cases.

Spatial Autoregressive Models (Part II)

District	Point estimate (Share SC)	Standard error	t-value
Jalaun	-0.21	0.06	-3.39
Jaunpur	-0.12	0.03	-3.93
Jhansi	-0.33	0.11	-2.95
Jyotiba Phule Nagar	-0.29	0.08	-3.75
Kannauj	-0.41	0.10	-4.08
Kanpur Dehat	-0.35	0.09	-4.12
Kanpur Nagar	-0.18	0.07	-2.71
Kaushambi	-0.37	0.09	-3.97
Kheri	-0.03	0.05	-0.71
Kushinagar	-0.23	0.08	-2.95
Lalitpur	-0.23	0.12	-1.87
Lucknow	0.05	0.04	1.12
Mahoba	-0.21	0.15	-1.38
Mahrajganj	-0.23	0.10	-2.38
Mainpuri	-0.42	0.09	-4.49
Mathura	-0.04	0.05	-0.82
Mau	-0.11	0.05	-2.29
Meerut	-	-	-
Mirzapur	0.00	0.05	0.02
Moradabad	-0.26	0.04	-6.35
Muzaffarnagar	-	-	-
Pilibhit	-0.24	0.07	-3.56
Pratapgarh	-0.17	0.06	-2.92
Rae Bareli	-0.06	0.04	-1.63
Rampur	-0.23	0.07	-3.51
Saharanpur	-	-	-
Sant Kabir Nagar	-0.16	0.06	-2.72
Sant Ravidas Nagar	-0.03	0.05	-0.74
Shahjahanpur	-0.34	0.04	-8.73
Shrawasti	-0.51	0.13	-4.03
Siddharthnagar	-0.34	0.06	-5.50
Sitapur	-0.21	0.05	-4.12
Sonbhadra	-0.31	0.05	-6.89
Sultanpur	-0.09	0.05	-1.65
Unnao	-0.18	0.06	-3.02
Varanasi	-	-	-

Table A43: Estimates from a spatial autoregressive model (part II). Estimates computed state-by-state due to the size of the spatial correlation matrix. Models could not converge in a few cases.

A10 ACCESS Survey

A10.1 Summary Statistics

- Table A44 provides the summary statistics for the ACCESS data used in the analysis.
- Table A45 reports the estimates of SC status on knowledge about RGGVY and household electrification.

	Mean	S.D.	Min.	Max.	Obs.
Heard of RGGVY	0.24	0.43	0	1	3023
Electrified (Grid)	0.57	0.49	0	1	3023
SC/ST	0.22	0.41	0	1	3023

Table A44: Summary statistics for ACCESS data used in the analysis.

A10.2 Evidence from Household Surveys

This section looks below the village level at the experiences of Dalit and non-Dalit households in Uttar Pradesh. We examine whether our community-level findings hold when we focus on individual households.

Collected between November 2014 and May 2015, the ACCESS survey data includes information from a representative sample of 252 villages from 21 districts (Aklin et al., 2016). The survey is useful because it contains questions about grid electrification status, awareness about RGGVY (i.e., whether the household head has heard of the scheme), and whether the household is Dalit or non-Dalit.

Table A45 uses the ACCESS survey to examine grid electricity connections, RGGVY awareness, and SC status. Models 1 and 4 are linear; models 2 and 4 are logistic regressions; models 3 and 6 are logistic regressions with conditional fixed effects. Standard errors are adjusted for sampling by village. As the table shows, SC households perform systematically worse than non-SC households. In model 1, we see that Dalit households are 4 percentage points less likely to have heard from the RGGVY, suggesting that RGGVY implementation is concentrated outside villages and habitations populated by Dalits. In model 4, we see that Dalit households are 15 percentage points less likely

	Heard of RGGVY			Electrified (Grid)		
	OLS	Logit	FE Logit	OLS	Logit	FE Logit
SC	-0.04**	-0.25**	-0.29**	-0.15***	-0.60***	-0.76***
	(0.02)	(0.12)	(0.14)	(0.02)	(0.09)	(0.11)
N	3023	3023	2255	3023	3023	2711
Villages	252	252	188	252	252	226

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A45: SC households, awareness of RGGVY (models 1-2), and household grid electricity. Models 2 and 4 are logistic regressions with sampling weights; models 3 and 6 are logistic regressions with conditional fixed effects at the village level. Standard errors are adjusted for sampling by village.

to have grid electricity connections, again consistent with the notion that the lack of RGGVY implementation is hurting Dalit households. Indeed, because this pattern is robust at the household level, the unequal pattern cannot be attributed to ecological inference problems.

Supplementary Appendix: References

Aklin, Michaël, Chao-yo Cheng, Johannes Urpelainen, Karthik Ganesan, and Abhishek Jain. 2016. "Factors Affecting Household Satisfaction with Electricity Supply in Rural India." *Nature Energy* 1 (16170).

Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142 (2): 615-635.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698-714.