Do Namesakes Confuse Voters? Evidence from Six Decades of Indian Elections

Gaurav Sood*

November 16, 2026

Abstract

Indian ballots often list minor independents who share a name with a leading candidate, and parties are widely reported to plant them to siphon votes. We ask three questions: do voters confuse namesakes with the candidates they resemble, how many votes does the confusion move, and is it enough to change who wins? Using 570,241 Lok Sabha and assembly candidate records (1962–2026), we compare independents whose names resemble a leading party candidate with other independents in the same contest. Namesakes of leaders place +7.8 percentile points higher (SE 1.4) among the contest's independents and poll about 19% more votes than a typical rival independent, while independents who resemble a minor party candidate gain nothing detectable. The gain is small in votes: the median namesake polls 462 votes, and their mean vote count is no higher than other independents'. Counting every vote a runner-up's namesakes received, they exceeded the winning margin in 76 of 57,166 contests; counting only the votes attributable to the name, 12 (95% range 4–18). The 2015 rule adding candidate photographs to voting machines coincides with a smaller premium, but placebo break dates and a trend in the premium produce changes of the same size, so the data cannot attribute the decline to photographs. Bihar's panchayat elections show no namesake premium.

1 Introduction

Parties and campaign operatives in India are widely reported to recruit "dummy" candidates: citizens who share the name of a leading contender and file as independents, in the hope that some of the contender's voters mark the wrong line (Factly, 2019; The Diplomat, 2020). The Election Commission of India cited the practice when it required electronic voting machines (EVMs) and ballot papers to carry candidate photographs from 1 May 2015 (Business Standard, 2015). Individual races look dramatic. In Mahasamund (Lok Sabha, 2014), Chandu Lal Sahu of the BJP won by 1,217 votes while 7 independents with the same or a same-sounding name polled 62,934 votes between them. The pattern is not chance: in Lok Sabha contests, a pair of one leading (top-two) candidate and one minor candidate is 2.6 times as likely to share an exact name as a pair of two minor candidates (7.6 versus 2.9 per 10,000 pairs; in assemblies 1.8 times, 8.4 versus 4.8).

*gsood07@gmail.com. I thank Priyadarshi Amar for obtaining the Election Commission's statistical reports for the 2023–2026 assembly elections.

Planting namesakes only matters politically if voters are fooled, and if enough are fooled to change results. Mahasamund illustrates the difficulty: the namesakes polled a lot, yet the candidate they resembled still won, and some of those votes would have gone to minor candidates anyway. This paper separates three questions. First, are voters confused, that is, does resembling a leader earn an independent votes it would not otherwise get? Second, how many votes does that move? Third, is it enough to overturn results? We then ask whether the 2015 photograph rule reduced confusion, and find that the data cannot answer that question.

2 Data and Design

Data. We use every general-election candidate in the Trivedi Centre for Political Data's Lok Dhaba files (Agarwal et al., 2021), which end with the 2019 Lok Sabha and the 2023 assembly elections, extended with the Election Commission's results for the 2024 Lok Sabha and the assemblies from November 2023 to 2026, which we put into TCPD's column layout with TCPD's own rules for its derived columns (`https://github.com/in-rolls/eci_results`; all are from the Commission's official statistical reports): 570,241 candidate records after dropping by-elections and state-years whose names are printed mostly as initials. Names are compared at three levels: exact or spelling-equivalent (identical after folding spellings of one spoken name), differing only in initials (*K.N. Rao* and *Rao*), and look-alike (one name contains the other, or the given names are at least 0.7 similar after shared words are set aside). The first two together are *close names*.

Comparison. Planted namesakes are independents, and almost all of them poll little. Comparing them with every losing candidate would set them against runners-up who poll 30–40% of the vote, so we compare independents with independents in the same contest. The sample is the 246,894 independents placed third or lower in 41,495 contests with at least two such independents; 17 of the 3,280 independents who resemble a leader finished first or second, so this restriction drops almost no namesakes while removing strong rebel independents from the comparison group. Table 2 column 4 keeps all independents. An independent is treated if its name resembles a party candidate who finished first or second (the *leader*). Because finishing first or second is itself an outcome, we also define the target before the election, as an incumbent or a national-party candidate (Table 2 column 3). TCPD identifies incumbents by linking candidates across elections, which the post-2023 results lack, so for every year we flag as incumbent the one candidate in each contest whose name best matches the winner of the seat's previous general election (token-set similarity of at least 90, seat boundaries unchanged, ties going to the winner's party). On TCPD's 2014–2023 rows this agrees with TCPD's flag for 98.6% of candidates; it finds 85% of TCPD's incumbents, and 91% of those it flags are incumbents by TCPD's flag. National parties are TCPD's through 2023 and, from 2024, the six recognized by the Election Commission in April 2023 (BJP, INC, BSP, CPI(M), AAP, NPP). This gives 245 pre-defined namesakes in 2024–2026, against 394 namesakes of leaders.

Outcomes. Each answers a different question. The candidate's *rank* among the contest's independents, scaled so the average independent sits at the 50th percentile, asks whether the name

helps at all and ignores how large any vote total is. log(1+votes) measures the proportional gain of a typical namesake. Votes estimated by Poisson regression measure the proportional gain in *mean* votes; the mean is dominated by the few independents who poll thousands, so this is the outcome that most depends on who else is in the comparison group.

Estimator. For independent i in contest c,

$$Y_{ic} = \alpha_c + \sum_k \beta_k D_{ic}^{(k)} + \sum_k \delta_k P_{ic}^{(k)} + \gamma \log(1 + \text{Freq}_i) + \varepsilon_{ic}, \tag{1}$$

where α_c is a contest fixed effect, so every comparison is between independents on the same ballot. $D_{ic}^{(k)}$ marks resemblance to a leader at level k, and $P_{ic}^{(k)}$ marks resemblance to a party candidate who finished third or lower. That second set is a placebo: such a candidate's voters are too few for confusion to pay, but the candidate's name is drawn from the same local pool of common names, so if common names attract votes on their own, δ_k will pick it up. Freq_i is how often the candidate's spelling appears among all candidates. Standard errors are clustered by constituency, because a constituency hosts many contests over six decades. Estimates come from `pyfixest`.

3 Are Voters Confused?

Table 1 gives the answer. Independents whose names match a leader's exactly or by spelling place +7.8 percentile points (SE 1.4) above the average independent on their ballot, and those differing only in initials +8.1 points (SE 1.2). In log votes, an exact namesake polls +0.17 (SE 0.04), about 19% more than a typical rival independent. The placebo shows that a common name alone does not do this: an independent who exactly matches a minor party candidate places +1.3 points higher (SE 1.7) and polls -0.02 log points (SE 0.05). Look-alikes are different. Their premium of +4.6 points (SE 0.9) is about twice the placebo's +2.3 (SE 0.9), so roughly half of the look-alike premium could be name commonness rather than confusion.

The premium does not depend on how the target is defined (Table 2). The close-name rank premium is +8.8 (SE 1.3) for namesakes of the winner and +7.1 (SE 1.1) for namesakes of incumbents and national-party candidates, defined before the vote. It is larger in Lok Sabha contests (+0.28 log points, SE 0.07) than in assembly contests (+0.14, SE 0.03), consistent with voters knowing less about candidates in constituencies of two million people.

4 How Many Votes?

Confusion is real but small in votes. The median leader namesake in our sample polls 462 votes. If the name multiplies a namesake's votes by e^{β}, the share owed to the name is $1 - e^{-\beta}$: 16% for exact namesakes and 15% for initials. For the median namesake that is a few dozen votes.

The Poisson estimates in the third column of Table 1 put an upper limit on this. A namesake's *mean* votes are no higher than other independents' (exact: -0.04, SE 0.12; initials: -0.11, SE 0.07), and when independents who finished first or second are included they are lower (-0.37, SE 0.09;

Table 1: Do independents who resemble a leader poll more than other independents on the same ballot?

	Rank (percentile pts.)	Log(1 + votes)	Votes (Poisson)
	(1)	(2)	(3)
Exact / spelling	7.78	0.17	-0.04
	(1.43)	(0.04)	(0.12)
Initials differ	8.06	0.16	-0.11
	(1.24)	(0.03)	(0.07)
Look-alike	4.65	0.09	-0.19
	(0.94)	(0.03)	(0.06)
Placebo: exact / spelling	1.26	-0.02	-0.02
	(1.68)	(0.05)	(0.15)
Placebo: initials differ	1.25	0.03	-0.02
	(1.95)	(0.05)	(0.16)
Placebo: look-alike	2.30	0.07	0.06
	(0.87)	(0.03)	(0.07)
Log name frequency	-0.19	-0.01	-0.01
	(0.06)	(0.00)	(0.00)
Contest fixed effects	x	x	x
Independent namesakes of a leader	3,148	3,148	3,148
Contests	41,495	41,495	41,495
Candidates	246,894	246,894	246,894

Independents placed third or lower, in contests with at least two of them. Rows 1-3: resemblance to a party candidate placed first or second; rows 4-6: to a party candidate placed third or lower. Standard errors clustered by constituency in parentheses.

Table 2: The namesake premium across targets and samples.

A. Rank among the contest's independents (percentile points)

	Top-2 leader	Winner	Incumbent or national party	All independents	Lok Sabha	Assembly
	(1)	(2)	(3)	(4)	(5)	(6)
Close name	7.93	8.82	7.09	7.53	11.26	7.14
	(0.96)	(1.35)	(1.13)	(0.96)	(2.54)	(1.02)
Look-alike	4.67	4.26	3.96	4.43	9.17	3.49
	(0.94)	(1.21)	(0.94)	(0.94)	(2.03)	(1.06)
Candidates	246,894	246,894	246,894	254,989	48,728	198,166

B. Log(1 + votes)

	Top-2 leader	Winner	Incumbent or national party	All independents	Lok Sabha	Assembly
	(1)	(2)	(3)	(4)	(5)	(6)
Close name	0.17	0.17	0.14	0.13	0.28	0.14
	(0.03)	(0.03)	(0.03)	(0.03)	(0.07)	(0.03)
Look-alike	0.09	0.06	0.08	0.07	0.23	0.05
	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)	(0.03)
Candidates	246,894	246,894	246,894	254,989	48,728	198,166

C. Votes (Poisson)

	Top-2 leader	Winner	Incumbent or national party	All independents	Lok Sabha	Assembly
	(1)	(2)	(3)	(4)	(5)	(6)
Close name	-0.07	-0.11	-0.09	-0.37	0.06	-0.22
	(0.07)	(0.08)	(0.07)	(0.09)	(0.12)	(0.06)
Look-alike	-0.19	-0.19	-0.15	-0.29	0.00	-0.36
	(0.06)	(0.07)	(0.06)	(0.08)	(0.08)	(0.08)
Candidates	246,894	246,894	246,894	254,989	48,728	198,166

Note: Close name is exact, spelling or initials resemblance; models also include log name frequency and contest fixed effects. Columns 1–3 vary the target: party candidates placed first or second, the winner, or incumbents and national-party candidates. Column 4 adds independents placed first or second. Standard errors clustered by constituency.

Table 2, column 4). Namesakes rise above the typical independent, but they do not join the few independents with a real following, who account for most independent votes. Read together, the outcomes say that resemblance shifts a namesake's vote count proportionally, from very little to a little more.

5 Is It Enough to Overturn Elections?

A runner-up's namesake can have cost the runner-up the seat only if the votes the name attracted exceed the winning margin. Table 3 counts such contests three ways. As an accounting upper bound, which treats every vote cast for the namesake as a vote meant for the runner-up, the runner-up's independent look-alikes polled more than the margin in 76 of the 1,368 contests where the runner-up had one. That bound means little on its own: the third-placed candidate's votes exceed the margin in 45% of all contests. Counting only the votes owed to the name, using each tier's log premium, the number falls to 12 contests (95% range 4–18), or 0.02% of the 57,166 contests with a party runner-up. For exact namesakes it is 1 (range 0–5). The Poisson estimates imply fewer still.

Two assumptions make these counts generous. Every vote owed to the name is assumed to have been meant for the runner-up rather than for the winner, another minor candidate, or abstention. And the proportional premium estimated for the typical namesake is assumed to apply to large namesakes too. Winners' namesakes cannot have changed results, since the candidate they resembled won.

Table 3: Contests in which a runner-up's independent namesakes polled more than the winning margin.

		Namesake votes exceed the margin		
	Runner-up has one	All votes (bound)	Votes owed to name	95% range
Exact / spelling	317	15	1	0-5
Close name (adds initials)	690	34	5	1-10
Any look-alike	1,368	76	12	4-18

Counts of contests among 57,166 general-election contests with a party runner-up. Votes owed to the name: each runner-up namesake's votes times 1 - exp(-b), b its tier's log premium in Table 1 (95% range from b's confidence interval).

6 Did Candidate Photographs Reduce Confusion?

The photograph rule is a natural test of the confusion account: once voters can see faces, a shared name should matter less. Comparing the premium before and after May 2015, it fell by -0.11 log points (SE 0.05; Table 4, column 4) and -3.6 rank points (SE 2.0). The data do not support attributing that decline to the photographs, for three reasons.

First, the premium was unstable long before 2015 (Figure 1, top). In log votes it was +0.28 in 1962–89 and +0.29 in the 1990s, fell to +0.07 in 2000–04 and +0.11 in 2005–09, and rose to +0.26

in 2010–April 2015. The post-rule values, +0.08 in May 2015–19 and +0.11 in 2020–26, are no lower than 2000–09. Second, a break placed at other dates produces changes of the same size (Figure 1, bottom). The real date gives $t = -2.3$; a break in 1998 gives $t = -2.8$, and 3 of 15 placebo dates reach $|t| > 1.96$. Third, allowing the premium to trend linearly leaves a change at May 2015 of -0.09 (SE 0.08), and restricting to 2005–2019 gives -0.11 (SE 0.07). Neither is distinguishable from zero.

Two more reasons for caution. The rule applied nationwide on one date, so the before-and-after comparison is one event, however many contests contribute to it. And the rule may have changed who runs as a namesake: if parties stopped planting them, the remaining namesakes are a different group.

Table 4: The namesake premium before and after the May 2015 photograph rule.

	Rank (percentile pts.)			Log(1 + votes)		
	1962-2026	+ trend	2005-2019	1962-2026	+ trend	2005-2019
	(1)	(2)	(3)	(4)	(5)	(6)
Close name	9.41	9.83	9.26	0.21	0.19	0.19
	(1.22)	(2.06)	(1.85)	(0.03)	(0.05)	(0.05)
Close name x after May 2015	-3.61	-4.26	-1.85	-0.11	-0.09	-0.11
	(1.97)	(3.05)	(3.06)	(0.05)	(0.08)	(0.07)
Close name x decades from May 2015		0.33			-0.01	
		(1.08)			(0.03)	
Look-alike	5.32	2.57	1.88	0.11	0.03	0.01
	(1.10)	(2.18)	(2.06)	(0.03)	(0.06)	(0.05)
Look-alike x after May 2015	-2.27	1.49	-2.89	-0.07	0.03	-0.06
	(2.05)	(3.25)	(3.50)	(0.05)	(0.09)	(0.09)
Look-alike x decades from May 2015		-1.57			-0.04	
		(1.03)			(0.03)	
Contest fixed effects	x	x	x	x	x	x
Candidates	246,894	246,894	57,629	246,894	246,894	57,629

Sample as in Table 1. Trend terms are in decades from 1 May 2015. Standard errors clustered by constituency in parentheses.

7 Local Elections: Bihar's Panchayats

Bihar's 2016 and 2021 panchayat elections cover 646,399 candidates placed third or lower in 155,194 contests across all six offices, from ward member to zila parishad. Candidates run without party labels, so the comparison is among all candidates placed third or lower. These elections are run by the State Election Commission, not the Election Commission of India, and we have not established whether their ballots carried photographs. Candidates who share the winner's name show no premium (Table 5): -0.5 rank points (SE 0.9), -0.01 log points (SE 0.02). In wards of a few hundred voters, who know the candidates personally, a shared name appears not to mislead.

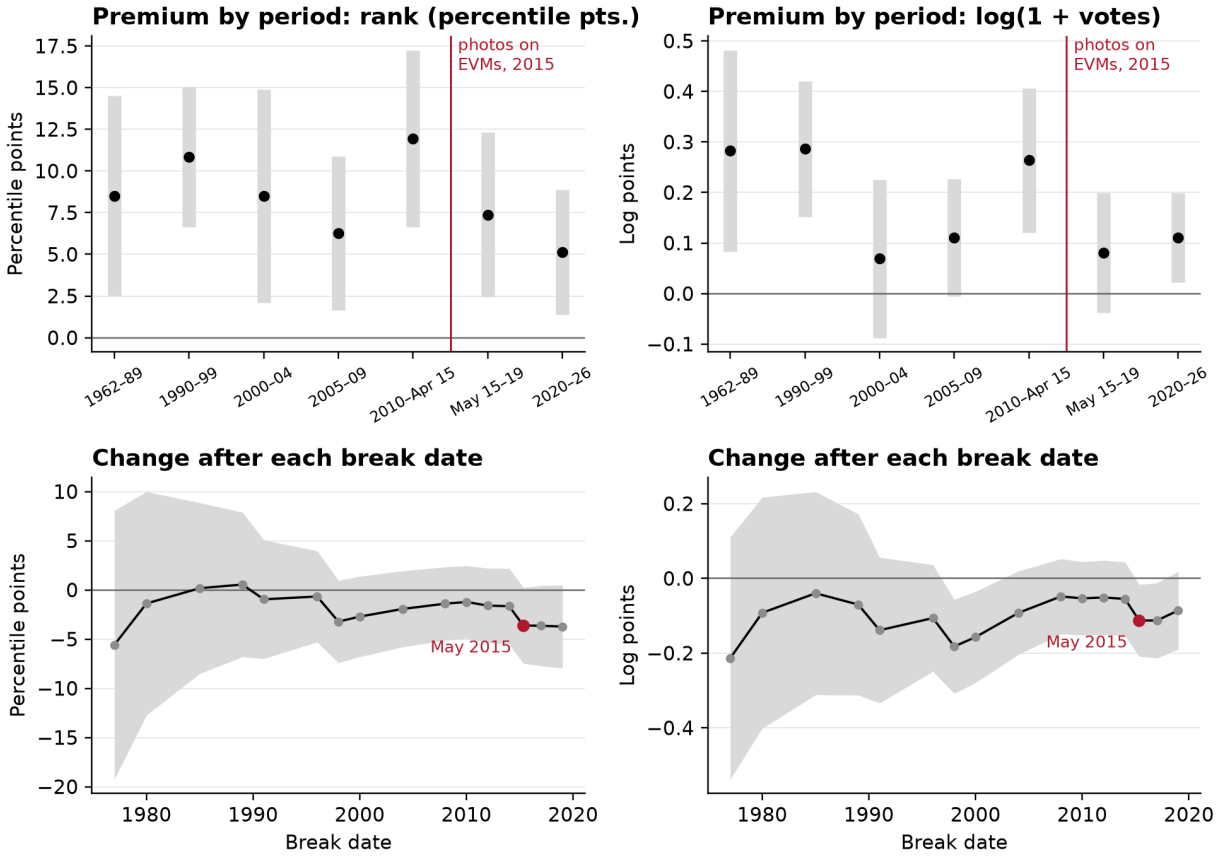


Figure 1: Top: close-name premium by period; the red line marks 1 May 2015. Bottom: the estimated change in the premium after each candidate break date, from a model with one break (Table 4, columns 1 and 4); gray dots are placebo dates and red marks the photograph rule. Gray bars and bands are 95% confidence intervals. Standard errors clustered by constituency.

Table 5: Bihar panchayat elections (2016, 2021): candidates who resemble the winner, against others placed third or lower.

	Rank (percentile pts.)	Log(1 + votes)	Votes (Poisson)
	(1)	(2)	(3)
Same name / alias	-0.47	-0.01	-0.02
	(0.88)	(0.02)	(0.03)
One name contains the other	-5.31	-0.09	-0.03
	(1.76)	(0.04)	(0.07)
Similarity 0.8+	-0.91	-0.01	-0.03
	(1.50)	(0.03)	(0.05)
Similarity 0.6-0.8	0.09	-0.00	-0.01
	(0.50)	(0.01)	(0.02)
Contest fixed effects	x	x	x
Contests	155,194	155,194	155,194
Candidates	646,399	646,399	646,399

Bihar panchayat candidates (2016, 2021, all six offices) placed third or lower, in contests with complete vote counts and at least two such candidates. Resemblance is to the contest's winner. All models control for log name frequency. Standard errors clustered by contest.

8 Conclusion

Voters are confused by namesakes: independents who resemble a leading candidate consistently outpoll other independents on the same ballot, and independents who resemble a minor candidate do not. The confusion moves few votes. Namesakes rise from very few votes to slightly more, and even counting generously, the votes owed to a runner-up's namesakes exceeded the winning margin in about 12 of 57,166 contests over six decades. Adding up every vote a namesake received, as press accounts do, gives 76 such contests instead of 12. Whether the 2015 photographs reduced confusion cannot be settled with these data, because the premium moved by as much at other dates.

References

Agarwal, Ananay, Saloni Agrawal, Saloni Bhogale, Sudheendra Hangal, Francesca R. Jensenius, Mohit Kumar, Shreyas Nizar and Gilles Verniers. 2021. "TCPD Indian Elections Data v2.0." Trivedi Centre for Political Data, Ashoka University.

Business Standard. 2015. "Ballot Papers, EVMs to Carry Candidates' Photographs from May 1." Press Trust of India / Business Standard, March 2015.

Factly. 2019. "Analysis: Do Dummy Candidates with Similar Names Affect Election Results?" `https://factly.in/analysis-do-dummy-candidates-with-similar-names-affect-election-results/`.

The Diplomat. 2020. "India's Fight Against Copying Candidates." https://thediplomat.com/2020/12/indias-fight-against-copying-candidates/.