Independent Candidates and
Political Representation in India

Online Appendix
(Not for Publication)

March 29, 2018

Online Appendix OA.1 Window Around the Reform

We estimate the effects of independents on voter turnout and representation in the 1996-1998 period around the reform. The estimates let us investigate whether our conclusions hold true in the smallest window of data possible.

Table A2 summarizes our main variables and estimates for the 1996 and 1998 elections. Like Table 2, it shows there were fewer independents across all constituencies, relatively fewer in open constituencies, and that independents are responsible for most of the reduction in the number of candidates overall. The number of independents decreases by 18.69 in open constituencies (Column 3) and by 10.51 in reserved constituencies (Column 6). The difference-in-differences estimate in Column 7 shows a relative decrease in open constituencies of 8.18 independents. The relative decrease is more pronounced than the one in the 1977-2004 sample (Table 2).

Like Table 2, Table A2 establishes a relative decline in voter turnout for open constituencies. Open constituencies have 3.52 percentage points more turnout after the reform. Reserved constituencies have 4.98 percentage points more. The differences-in-differences estimate shows a relative decrease of 1.46 percentage points in open constituencies, less pronounced than the estimated effect we obtain using the full 1977-2004 sample.

Statistics and estimates for our measure of representation in government are found in the third to last row of Table A2. This row shows that the probability of electing a governing-coalition candidate increased by more in open constituencies. The increases in open and reserved constituencies were 22 and 7 percentage points, respectively. The relative increase of 15 percentage points is more pronounced than the relative increase for the 1997-2004 sample.

The corresponding IV estimates are found in the third and fifth rows of Table A2. The IV estimate for turnout is the ratio of its reduced-form estimate (second row) to the first-stage

estimate, multiplied by the standard deviation for the number of independents, which equals 22.91 candidates in the present sample. The IV estimate for the probability of electing a governing-coalition candidate is defined similarly.

The third row shows that a one standard deviation increase in the number of independents increases voter turnout by 4.10 percentage points. The estimate is approximately 2 percentage point smaller than our most flexible estimate for the 1977-2004 sample (Column 5 of Table 4). The estimate is less precise, however, as it has a p-value of 0.11. The fifth row shows that a one standard deviation increase in the number of independents decreases the probability of electing a governing-coalition candidate by 42 percentage points. The estimate is about 11 percentage points larger than our most flexible estimate for the 1977-2004 sample (Column 5 of Table 7). Although they differ in precision and magnitude, the estimates support the qualitative conclusions drawn from the 1977-2004 sample.

It is important to note that a shorter panel of elections is not without disadvantages. Figure OA1(a) illustrates why. The figure plots $\Delta Turnout_j = Turnout_{j98} - Turnout_{j96}$ and $\Delta(\#Independents)_j = \#Independents_{j98} - \#Independents_{j96}$ for all 543 constituencies. It shows extreme changes in turnout and independents in several constituencies between 1996 and 1998, and emphasizes the dominance of these changes over changes in other constituencies. The figure suggests that to obtain credible difference-in-differences estimates with a shorter panel we should either drop the extreme observations or consider an alternative to a least squares estimator.

Figure OA1(b) plots $\Delta Turnout_j$ and $\Delta(\#Independents)_j$ for the 532 constituencies where changes in turnout and the number of independents were relatively moderate. These are constituencies where the number of independents declined by less than 60, or where turnout changed by 30 percentage points or less.[1] The figure excludes constituencies which, arguably, are not part of the population of interest. For example, it excludes constituencies where there were large-scale boycotts of the 1998 election (and consequently large declines

---

[1]The excluded constituencies all had abnormally large values for cook's distance measure (of influence) in either the first stage and the reduced form of our statistical model.

in turnout), constituencies like Nagaland, Outer Manipur, and several in the state of Assam (**?**).[2]

Table OA1 replicates Table A2 using the 532 constituencies described in Figure OA1(b). It shows that a one standard deviation increase in the number of independents (11.83 here) increases voter turnout by 3.25 percentage points and decreases the probability of electing a governing-coalition candidate by 33 percentage points. Here both coefficients are precisely estimated at conventional significance levels (p is 0.08 and 0.00, respectively). The estimates from this no-outlier sample further support our qualitative conclusions from the 1977-2004 sample.

---

[2]Figure OA1(b) shows one constituency (Baramulla) where the number of independents increased substantially following the reform. The increase is emblematic of the volatility of elections in Jammu and Kashmir.
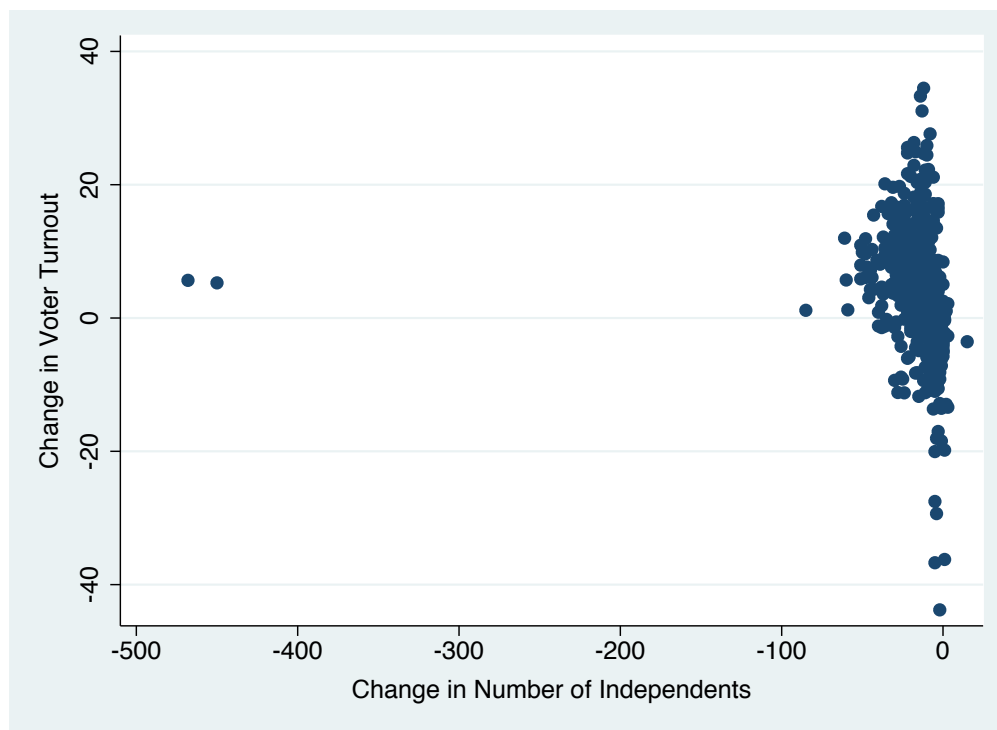
(a) With Influential Observations (N=543)
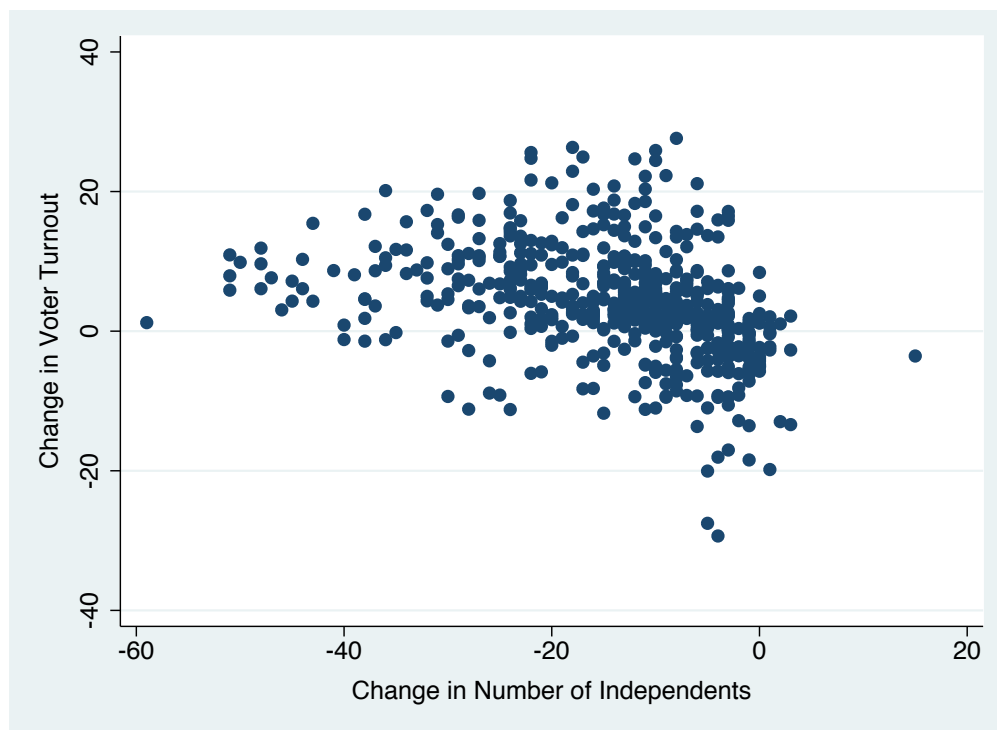
(b) Without Influential Observations (N=532)

Figure OA1: Changes in Voter Turnout and the Number of Independents with and without Influential Observations.

Table OA1: Descriptives for 1996 and 1998 Elections Alone. Standard deviations for means are in round parentheses. Standard errors (in square parentheses) for mean comparisons are clustered at the level of the constituency.

	Open Constituencies (Seat can be held by anyone)			Reserved Constituencies (Seat can only be held by Minorities)			Differences-in-Differences
	Minorities pay 250 Rupees, Majorities pay 500 Rupees	Minorities pay 5000 Rupees, Majorities pay 10000 Rupees	Difference	Minorities pay 250 Rupees, Majorities pay 500 Rupees	Minorities pay 5000 Rupees, Majorities pay 10000 Rupees	Difference	
	(1)	(2)	(2) - (1)	(4)	(5)	(5)-(4)	(6) - (3)
# Independent Candidates (Independents or 1-Member Unrecognised)	20.09 (13.52)	4.13 (3.71)	-15.96 [0.63]	13.19 (10.20)	2.67 (2.42)	-10.52 [0.69]	-5.44 [0.93]
Voter Turnout (% Eligible Electors)	57.96 (12.37)	61.53 (9.24)	3.56 [0.45]	56.23 (13.63)	61.29 (9.08)	5.06 [0.57]	-1.50 [0.72]
IV Estimate of a one-standard deviation increase in the Number of Independents on Voter Turnout							3.25 [1.86]
Elected Representative is a Member of Governing Coalition	0.27 (0.44)	0.49 (0.50)	0.22 [0.03]	0.35 (0.48)	0.42 (0.49)	0.07 [0.05]	0.15 [0.05]
IV Estimate of a one-standard deviation increase in the Number of Independents on Member of Government							-0.33 [0.14]
Elections	1	1	2	1	1	2	2
Constituencies	360	360	360	172	172	172	532
Observations	360	360	720	172	172	344	1064

5

Online Appendix OA.2 Dummy Candidates

Estimates are found in Table OA2. The first column presents the effect on the average Q-gram string distance between the candidates of an election. This measure takes all the 3-gram vectors of a pair of strings and counts the number of differences. For example, Arvind and Arvint have 3-grams of 'Arv', 'rvi', 'vin', 'ind', 'int', and a distance of 2. The second column presents the effect on the Jaro-Winkler string distance. The measure lies between between 0 and 1, where 0 implies candidate names are not at all similar, and 1 implies the similarity is exact. The third column presents the effect on the the Levenhstein distance. This measure counts the number of deletions, insertions, and substitutions it takes to go from the name of one candidate to the name of another. The fourth column presents the the longest common substring distance. This measure counts the minimum number of characters one must remove (from any of pair of names) before getting the same substring (Dist(SAchA, SArA) = 3). Note that we will multiply the Jaro-Winkler measure by -1. Because of this, for all the measures, the larger its value, the more dissimilar are the candidate names.

If anything the estimates imply higher deposits led to candidate names that were more similar. The higher deposits decreased the average Q-gram distance by 0.08. It decreased the average least common substring distance by 0.07. It decreased the average Levenhstein distance by -0.04, though the estimate is statistically insignificant at conventional levels. The average for Jaro-Winkler distance measure was unaffected by the higher entry deposits. Ultimately all four estimates imply the effect on the presence of fake candidates was small to negligible.[3]

---

[3]While cleaning the data we looked for the use of fake candidates. We found some instances where it was apparent that this was going one. But these instances were few and far between.

Table OA2: Dummy Candidates. This table examines whether low entry deposits encouraged the use of fake candidates. Fake candidates will have names that are similar to the names of the more serious candidates. Fake candidates can confuse voters and decrease the vote shares of these serious candidates. The dependent variables are different measures of the average (string) distance over all the pairwise names of the candidates who participated in an election. The dependent variables are all increasing in the dissimilarity of candidate names. Fake candidates are more likely where candidate names are most similar. LCS is the acronym for least common substring. Standard errors (in parentheses) are clustered at the level of constituency.

	Measure of Dissimilarity of Candidate Names			
	Q-gram	Jaro-Winkler	Levenh-stein	LCS
	(1)	(2)	(3)	(4)
Open Seat After the Fee Increase	-0.08	-0.00	-0.04	-0.07
	(0.11)	(0.00)	(0.06)	(0.08)
Mean of Dependent Variable	24.50	0.47	14.04	19.48
Constituency Fixed Effects	✓	✓	✓	✓
Election Year Fixed Effects	✓	✓	✓	✓
Independents in 1977 (Trends)	✓	✓	✓	✓
Dependent Variable in 1977 (Trends)	✓	✓	✓	✓
Eligible Electors (in logs)	✓	✓	✓	✓
Distance Between Winner and and Runner-up in Last Election	✓	✓	✓	✓
Observations	4282	4297	4297	4297

Online Appendix OA.3 Preferences and Information

We investigate the robustness of our main results to differences in the evolution of voter preferences across open and reserved constituencies. Our investigation draws on data from the 1971 Indian National Election Study. The data has information on voter perceptions, including their perceptions about the effectiveness of their local government, about whether they feel pressure to vote a certain way. It has information on whether they actively pursued knowledge about the election (via newspapers or radio), whether they have a specific interest in national or state politics, whether they have a general interest and discuss politics, and the perceived complexity of politics and government. It has a measure of their preference for redistributive government policies. In all the data is based on interviews with 640 voters, from 78 constituencies, of which 19 have seats that reserved for disadvantaged persons.[4]

We include interactions of these measures with a fourth-order polynomial in time in our reduced form and first stage specifications. Estimates are found in Table OA3. Moving left to right shows how the coefficients of interest change as the interactions are included. We exclude the IV estimates because, with this small sample, the first stage is not strong enough.

Our main (reduced form) results are similar even though the data only includes a small cross-section of constituencies from our sample. In these constituencies, the reduction in independent candidates ranges from between 1.48 and 2.96 candidates. The reduction in voter turnout ranges from 0.64 to 2.38 percentage points. The probability that the elected representative is a member of the governing coalition increases by between 0.07 and 0.09 percentage points. At least for this small sample of constituencies, the estimates of interest are not explained away by systematic evolutionary differences across open and reserved constituencies. We find this unsurprising in large part because of the difficulty with finding a differential trend that explains the sizeable and sudden drop in the number of independents.

---

[4]For more details about the questions and methodology see http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/25402.

Table OA3: Voter Preference and Information.

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable = # Independents					
Open Seat After the Fee Increase	-1.81	-2.14	-3.17	-2.77	-2.70	-3.17
	(1.10)	(1.10)	(0.94)	(1.04)	(1.13)	(1.21)
	Dependent Variable = Voter Turnout					
Open Seat After the Fee Increase	-3.02	-2.87	-2.16	-1.40	-1.62	-1.49
	(1.11)	(1.06)	(0.96)	(1.11)	(1.12)	(1.15)
	Dependent Variable = Elected Representative is a Member of the Governing Coalition					
Open Seat After the Fee Increase	0.11	0.12	0.10	0.08	0.08	0.08
	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
(National) Constituencies	78	78	78	78	78	78
Observations	569	569	569	569	569	569
Local Government is Attentive	✓	✓	✓	✓	✓	✓
Villagers Influence Votes of Others	✓	✓	✓	✓	✓	✓
Get Election News from Newspapers or Radios		✓	✓	✓	✓	✓
Interested in National Politics, State Politics or Both			✓	✓	✓	✓
General Interest in Politics or Discusses Politics				✓	✓	✓
Politics and Government are too Complicated					✓	✓
Believe in Redistributive Government Policy						✓

9

Online Appendix OA.4 Miscellaneous Tables and Robustness Checks

Table OA4: Robustness to Alternative Definitions of Independent Candidates. Standard errors (in parentheses) are clustered at the level of constituency.

| | First Stage Estimates: # Independents | | | | | Instrumental Variables Estimates | | | |
	(1)	(2)	(3)	(4)	(5)	Voter Turnout (6)	Member of Government (7)	Vote Shares Independents (8)	Winner (9)
Election of Commission's Official Definition of Independents									
Open Seat After the Fee Increase	-2.64 (0.47)	-2.03 (0.46)	-2.04 (0.47)	-2.07 (0.49)	-2.61 (0.60)				
# Independents						0.49 (0.23)	-0.03 (0.01)	0.87 (0.30)	-0.38 (0.20)
Our Definition: Official Definition + Unrecognised Parties with One Candidate Running									
Open Seat After the Fee Increase	-2.62 (0.47)	-2.03 (0.46)	-2.04 (0.47)	-2.08 (0.49)	-2.63 (0.60)				
# Independents						0.49 (0.22)	-0.03 (0.01)	0.89 (0.31)	-0.38 (0.20)
Alternative Definition: Official Definition + All Candidates from Unrecognised Parties									
Open Seat After the Fee Increase	-2.59 (0.00)	-2.07 (0.00)	-2.05 (0.00)	-2.11 (0.00)	-2.66 (0.00)				
# Independents						0.49 (0.22)	-0.02 (0.01)	0.57 (0.40)	-0.36 (0.19)
Observations	4840	4840	4840	4840	4297	4297	4297	4026	4297
Constituency Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Election Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Independents in 1977 (Trends)		✓	✓	✓	✓	✓	✓	✓	✓
Dependent Variable in 1977 (Trends)			✓	✓	✓	✓	✓	✓	✓
Eligible Electors (in logs)				✓	✓	✓	✓	✓	✓
Distance Between Winner and Runner-up in Last Election					✓	✓	✓	✓	✓

10

Table OA5: Robustness to Exclusion of Constituencies with more than 100 Candidates. Standard errors (in parentheses) are clustered at the level of constituency.

	First Stage Estimates					Instrumental Variables Estimates			
	# Independents					Voter Turnout	Member of Government	Vote Shares	
								Independents	Winner
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Open Seat After the Fee Increase	-2.18	-1.54	-1.56	-1.59	-1.99				
	(0.37)	(0.34)	(0.34)	(0.36)	(0.42)				
# Independents						0.65	-0.033	1.20	-0.49
						(0.29)	(0.017)	(0.41)	(0.25)
Observations	4837	4837	4837	4837	4294	4294	4294	4037	4294
Constituency Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Election Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Independents in 1977 (Trends)		✓	✓	✓	✓	✓	✓	✓	✓
Dependent Variable in 1977 (Trends)			✓	✓	✓	✓	✓	✓	✓
Eligible Electors (in logs)				✓	✓	✓	✓	✓	✓
Distance Between Winner and Runner-up in Last Election					✓	✓	✓	✓	✓

11

Table OA6: Robustness to Exclusion of Post-1998 Elections. Standard errors (in parentheses) are clustered at the level of constituency.

| | First Stage Estimates | | | | | Instrumental Variables Estimates | | | |
| | # Independents | | | | | Voter Turnout | Member of Government | Vote Shares Independents | Vote Shares Winner |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Open Seat After the Fee Increase	-2.80	-1.94	-1.94	-2.03	-2.58				
	(0.48)	(0.49)	(0.51)	(0.55)	(0.66)				
# Independents						0.77	-0.04	0.88	-0.51
						(0.36)	(0.02)	(0.42)	(0.26)
Observations	3754	3754	3754	3754	3211	3211	3211	3067	3211
Constituency Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Election Year Fixed Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Independents in 1977 (Trends)		✓	✓	✓	✓	✓	✓	✓	✓
Dependent Variable in 1977 (Trends)			✓	✓	✓	✓	✓	✓	✓
Eligible Electors (in logs)				✓	✓	✓	✓	✓	✓
Distance Between Winner and Runner-up in Last Election					✓	✓	✓	✓	✓

Table OA7: Robustness to Differential Trends in Rural Population. Standard errors (in parentheses) are clustered at the level of constituency.

	Voter Turnout (% of Eligible Electors)				
	(1)	(2)	(3)	(4)	(5)
	Instrumental Variables				
# Independents	0.99	1.09	0.55	0.50	0.47
	(0.27)	(0.38)	(0.26)	(0.25)	(0.22)
$S-W$ F-Test of Excluded	30.49	18.47	18.80	18.15	19.28
Instruments, $F(1, 542)$	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Rural Population Share in 2001 (Trends)	✓	✓	✓	✓	✓
Constituency Fixed Effects	✓	✓	✓	✓	✓
Election Year Fixed Effects	✓	✓	✓	✓	✓
Independents in 1977 (Trends)		✓	✓	✓	✓
Turnout in 1977 (Trends)			✓	✓	✓
Eligible Electors (in logs)				✓	✓
Distance Between Winner and Runner-up in Last Election					✓
Observations	4807	4807	4807	4807	4268