Online supporting information for "How governments shape the risk of civil violence: India's federal reorganization, 1950–56"

Bethany Lacina
Assistant Professor,
Department of Political Science,
University of Rochester
blacina@ur.rochester.edu

1 Overview

This appendix contains summary statistics for variables used in the main analysis, a table and map of all language enclaves analyzed in the main text; a comparison of the *Linguistic Survey of India* (*LSI*) to the 1951 census for select Indian states; and checks of the robustness of the results in the main text.

2 Summary statistics for variables used in main text

Table 1: Summary statistics for variables used in main text

	Mean	St. Dev.	Min	Max
Violence	0.25	0.44	0	1
Peaceful statehood	0.14	0.35	0	1
Statehood	0.25	0.44	0	1
Ln relative INC representation	-1.4	1.7	-4.9	4.2
Ln relative INC representation sq.	5.0	5.3	0	24
Demographic polarization	0.051	0.046	0.00065	0.22
Cultural polarization	0.032	0.026	0.00028	0.12
Ln enclave plurality group's INC rep.	0.34	1.4	-2.0	3.4
Ln enclave plurality group's population	14	1.5	11	17
Agricultural labor share in enclave	0.74	0.19	0.012	0.98
Landless rate in enclave	0.13	0.091	0.00024	0.34
Hindu share in enclave	0.78	0.24	0.0057	0.99
Ln km to New Delhi	6.5	1.3	0.00066	8.9
Violence * Ln relative INC rep.	-0.27	0.84	-2.9	3.5
Observations	63			

3 Map and list of enclaves

Table 2 is a list of all language enclaves in my data and codes for locating the enclave on a map of India (Figure 1). Note that Jammu and Kashmir, the Northeast Frontier Administration (NEFA),

and the Andaman and Nicobar Islands were not fully under government control in 1951 and are not included in my data. Also, three groups listed in Table 2—Bihari in Bihar, Oriya in Orissa, and Bengali in West Bengal—were majority languages in Type A states in 1950. These groups are not included in my regression analyses but are in Table 2 and Figure 1 for reference. Table 2 also lists the values of the dependent variables of violence and accommodation. Cases that are coded as accommodated but not coded as violent are the instances of peaceful accommodation in my data.

Figure 1: Enclaves with potential statehood claims during India's reorganization

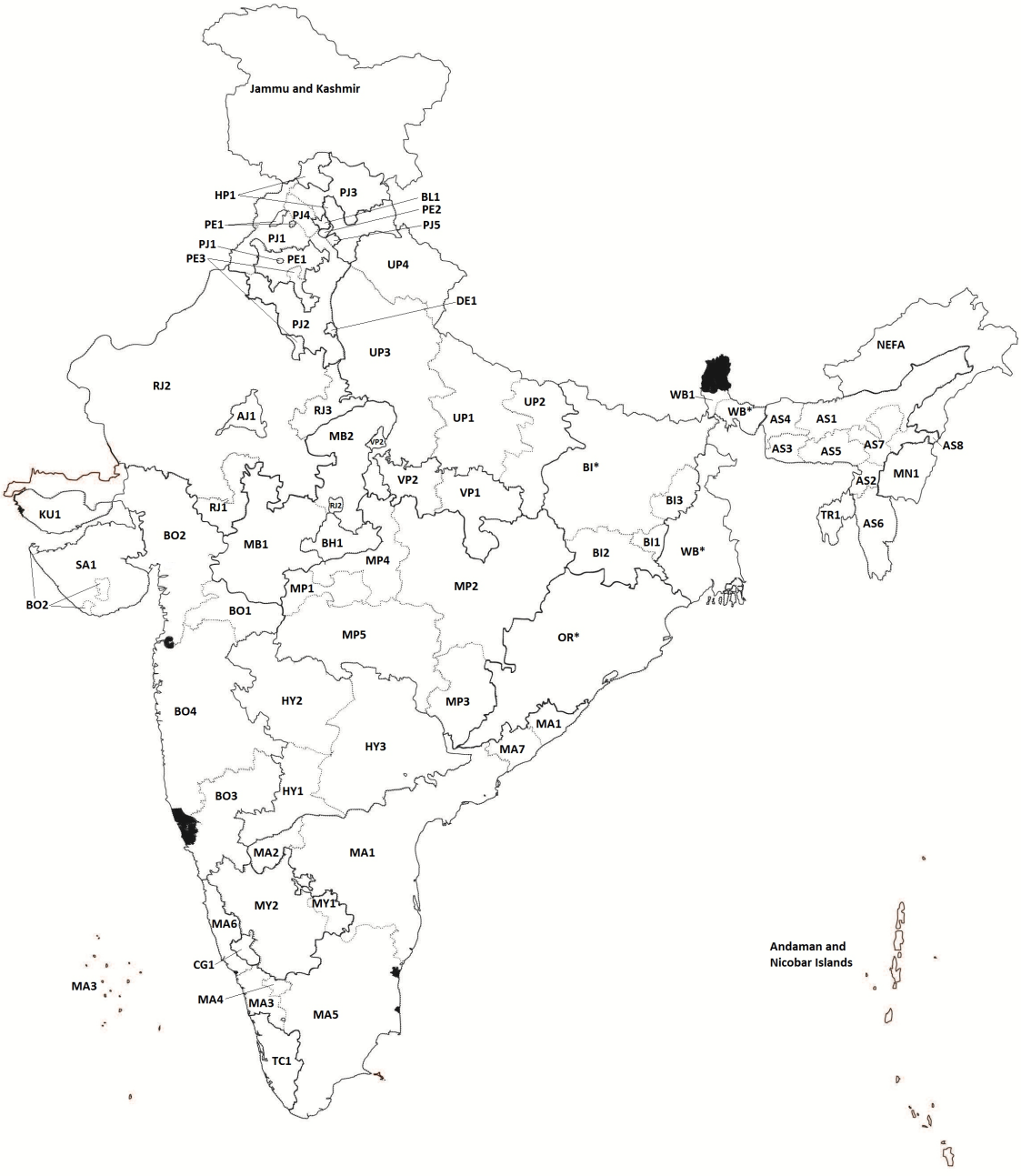


Table 2: Enclaves with potential statehood claims during India's reorganization

Enclave	State	Plurality language	Map	Violence?	Accommodation?
Ajmer	Ajmer	Rajasthani	AJ1		Y
Assamese–Assam	Assam	Assamese	AS1		
Cachar	Assam	Bengali	AS2		
Garo Hills	Assam	Garo	AS3		
Goalpara	Assam	Bengali	AS4	Y	
Khasi Hills	Assam	Khasi	AS5		
Lushei Hills	Assam	Mizo	AS6		
Mikir Hills	Assam	Karbi	AS7		
Naga Hills	Assam	Naga	AS8	Y	
Bhopal	Bhopal	Rajasthan	BH1		
Bihar	Bihar	Bihari	BI*		
Purulia	Bihar	Bengali	BI1	Y	Y
Ranchi and Singhbhum	Bihar	Santali	BI2		
Santal Parganas	Bihar	Santali	BI3		
Bilaspur	Bilaspur	Pahari	BL1		
The Dangs	Bombay	Khandesi	BO1		
Gujarat	Bombay	Gujarati	BO2	Y	
Kannada–Bombay	Bombay	Kannada	BO3	Y	Y
Maharashtra	Bombay	Marathi	BO4	Y	
Coorg	Coorg	Kannada	CG1		Y
Delhi	Delhi	W Hindi	DE1		
Himachal Pradesh	Himachal Pradesh	Pahari	HP1		
Karnataka	Hyderabad	Kannada	HY1		Y
Marathwada	Hyderabad	Marathi	HY2		
Telangana	Hyderabad	Telugu	HY3		Y
Kutch	Kutch	Sindhi	KU1		
Malwa	Madhya Bharat	Rajasthani	MB1		
Gwalior	Madhya Bharat	W Hindi	MB2		
Bhopawar	Madhya Pradesh	Rajasthani	MP1		
Chhattisgarh	Madhya Pradesh	E Hindi	MP2		
Gondh	Madhya Pradesh	Gondi	MP3		
Mahakoshal	Madhya Pradesh	W Hindi	MP4		
Vidarbha	Madhya Pradesh	Marathi	MP5		

Continued...

Enclave	State	Plurality language	Map	Violence?	Accommodation?
Andhra	Madras	Telugu	MA1	Y	Y
Bellary	Madras	Kannada	MA2	Y	Y
Malabar	Madras	Malayalam	MA3		Y
Nilgiris	Madras	Kannada	MA4		
Tamil Nadu	Madras	Tamil	MA5	Y	Y
Tulu Nad	Madras	Tulu	MA6		
Visagapatnam	Madras	Oriya	MA7	Y	
Manipur	Manipur	Meitei	MN1	Y	
Kolar	Mysore	Telugu	MY1		
Mysore	Mysore	Kannada	MY2	Y	Y
Orissa	Orissa	Oriya	OR*		
Gurumukhi–PEPSU	PEPSU	Gurumukhi	PE1	Y	
Punjabi–PEPSU	PEPSU	Punjabi	PE2		
Sangrur	PEPSU	W Hindi	PE3		
Gurumukhi–Punjab	Punjab	Gurumukhi	PJ1	Y	
Haryana	Punjab	W Hindi	PJ2		
Kangra	Punjab	Pahari	PJ3		
Punjabi–Punjab	Punjab	Punjabi	PJ4	Y	
Simla	Punjab	Pahari	PJ5		
Bhilwara	Rajasthan	Bhili	RJ1		
Rajasthani–Rajasthan	Rajasthan	Rajasthani	RJ2		Y
Swai Madhopur	Rajasthan	W Hindi	RJ3		
Saurashtra	Saurashtra	Gujarati	SA1		
Kerala	Travancore-Cochin	Malayalam	TC1		Y
Tripura	Tripura	Bengali	TR1		
Awadh	Uttar Pradesh	E Hindi	UP1		
Bhojpur	Uttar Pradesh	Bihari	UP2		
Harit Pradesh	Uttar Pradesh	W Hindi	UP3		
Tehri-Garhwal	Uttar Pradesh	Pahari	UP4		
Baghelkhand	Vindhya Pradesh	E Hindi	VP1	Y	
Bundelkhand	Vindhya Pradesh	W Hindi	VP2		
Bengal	West Bengal	Bengali	WB*		
Darjeeling	West Bengal	Nepali	WB1		

4 The *Linguistic Survey of India* and the 1951 census compared

This section investigates whether the *LSI* data is a good approximation of language composition of Indian districts in the 1950s, using three states where language categories are comparable in the *LSI* and the 1951 Indian census.

I selected three states where the language names used in the *LSI* and the 1951 census closely correspond: Bombay state, West Bengal, and Orissa. For each district in those three states, I compiled all the languages with at least 10% of the district population according to the *LSI*. I noted each language's district-level population share according to the *LSI*. I matched those languages to the district data from the 1951 census, recording the census' estimate of each language's district population share. If a language appeared in the *LSI* but not the census, the census estimate of the district population share of that language was coded as zero. In case of a language that was recorded in the census as comprising $\geq 10\%$ of the population of a district but was not in the *LSI*, I added an observation for that language, coding its *LSI* population share estimate as zero.

The resulting data are language-district format. The two variables are population shares per the linguistic survey, named *LSI share* ($\mu = 0.34$, $\sigma = 0.37$, range $= 0 - 0.99$), and population shares per the 1951 census, named *Census share* ($\mu = 0.32$, $\sigma = 0.39$, range $= 0 - 0.99$).

Figure 2 plots these two variables against each other along with a least-squares fit line. Table 3 reports the ordinary least-squares regression of one estimate on the other. The estimated regression coefficient is highly significant and close to one, indicating that the two sources correspond quite well.

The outliers in the analysis are several districts in Bombay state that the *LSI* records as majority tribal and the census records as majority Gujarati or Marathi. One explanation of this difference is linguistic assimilation of tribal populations between the 1890s and 1950s. However, during the 1951 census there was also violent intimidation of tribal populations by larger language groups looking to solidify their claims in anticipation of state reorganization. That kind of politicization

Figure 2: Scatterplot of *Linguistic Survey of India* and 1951 census estimates of district-level population shares for all languages in Bombay, Orissa, and West Bengal

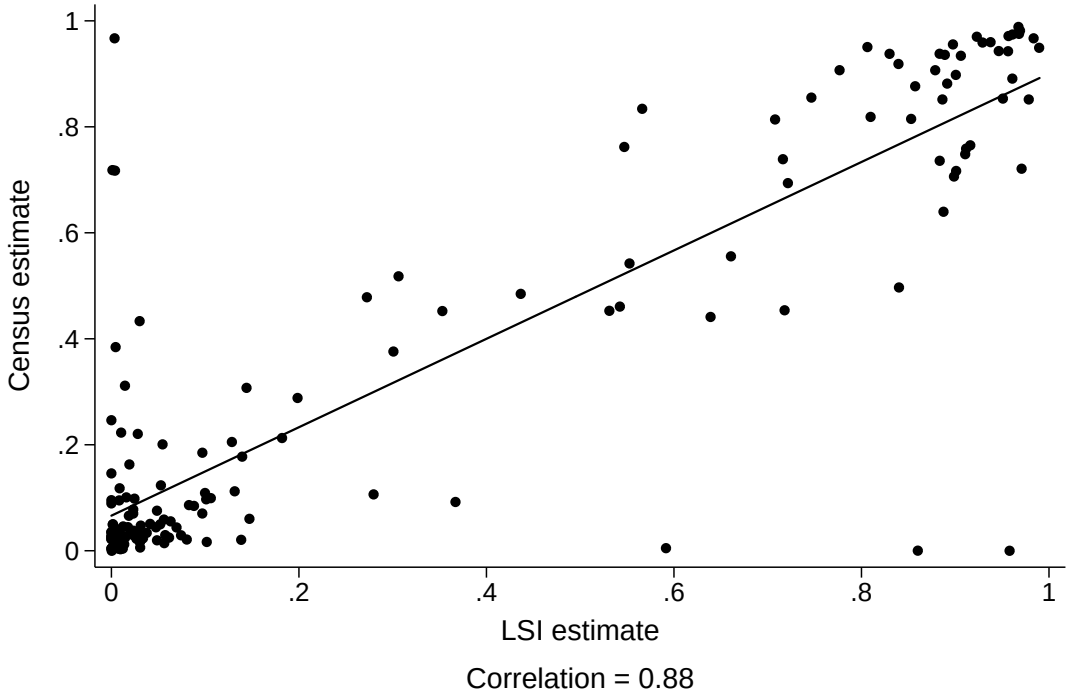


Table 3: OLS regression of *Linguistic Survey of India* and 1951 census estimates of district-level population shares for all languages in Bombay, Orissa, and West Bengal

	Model 1
	Census estimate
LSI estimate	0.92***
	(0.041)
Constant	0.015
	(0.020)
Observations	159
R-squared	0.77

Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

of the census suggests that the *LSI* may be the more accurate source in cases where the two sources diverge.

5 Robustness tests of regression results in main text

The remainder of the appendix presents a series of regressions that recreate the analyses in the main text, performing additional checks of robustness. Table 4 presents summary statistics for all the variables introduced. These variables are all used in robustness tests of the multinomial logistic regressions of peaceful accommodation and violence (Table 5 in the main text). Where theoretically relevant, these variables are also used in robustness tests of the logistic regressions of all accommodation (Table 6 in the main text).

5.1 Alternative measures of demographic and cultural polarization

In the main text, two controls are introduced for the demographic balance between the groups at odds over state reorganization. The main text measures cultural polarization as follows:

$$\textit{Cultural polarization} = (n_i^2 n_j + n_i n_j^2)(1 - s_{ij}^{\delta}) \tag{1}$$

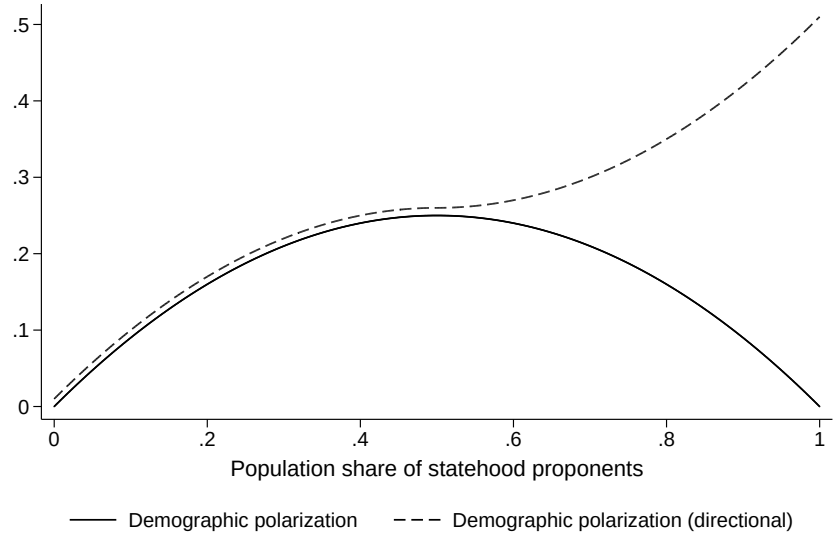
where s_{ij} is the number of common branches in a universal table of language genealogy, divided by 15, and $\delta = 0.5$. Esteban et al. (2012) use a lower value of δ (0.05), which has the effect of weighting cultural difference less heavily. *Cultural polarization II* is calculated from equation 1 and $\delta = 0.05$. This variable is included in a multinomial model for peaceful accommodation and violence (Table 5, Model 2). Results are quite similar to those in the main text.

It may be important to distinguish cases of polarization according to whether the proponents or opponents of statehood had the edge in population size. The distinction between low levels of polarization that favor the proponents of statehood versus low levels of polarization that disfavor

Table 4: Summary statistics for variables introduced in robustness tests

	Mean	St. Dev.	Min	Max
Cultural polarization II	0.015	0.026	0.000035	0.12
Demographic polarization (directional)	0.079	0.11	0.00065	0.49
Cultural polarization (directional)	0.060	0.11	0.00028	0.50
Enclave demographic polarization	0.22	0.27	3.7 e-6	1.00
Enclave cultural polarization	0.090	0.10	1.6 e-6	0.32
Enclave fractionalization	0.41	0.20	0.0013	0.79
Colonial military different	0.32	0.47	0	1
Colonial admin. different	0.43	0.50	0	1
Colonial military disadvantaged	0.13	0.34	0	1
Colonial admin. disadvantaged	0.14	0.35	0	1
Total claims in state	4.8	3.4	1	12
State demographic polarization	0.29	0.19	0.11	1.00
State cultural polarization	0.29	0.19	0.11	1.00
State fractionalization	0.61	0.18	0.0013	0.79
Agricultural labor share in state	0.74	0.12	0.12	0.94
Landless rate in state	0.15	0.077	0.0077	0.32
Inequality (State vs. enclave)	0.29	2.1	6.9 e-19	16
Relative deprivation (State vs. enclave)	0.58	0.56	0	1.3
Relative wealth (State vs. enclave)	1.4	7.2	0	58
Average winner's vote share in enclave	0.52	0.12	0	1
Effective parties in Vidhan Sabha	2.4	1.2	1.2	5.4
INC seat share in Vidhan Sabha	0.68	0.17	0.41	0.94
PCC since 1921	0.21	0.41	0	1
Ln km to Varanasi	6.8	0.59	5.1	8.9
Land or sea border	0.60	0.49	0	1
Regionalist party vote share in enclave	0.064	0.13	0	0.46
Hindu right vote share in enclave	0.068	0.088	0	0.38
Ln km to Bombay	6.9	0.66	5.0	9.0
Vernacular daily newspapers in enclave	5	9.2	0	55
English daily newspapers in enclave	1.3	4.1	0	26
Observations	63			

Figure 3: Sample calculations of demographic polarization and a directional measure of demographic polarization



the same group could be particularly relevant to the incidence of peaceful accommodation. I code a directional measure of demographic polarization:

$$\textit{Demographic polarization (directional)} = \begin{cases} n_i^2 n_j + n_i n_j^2 & \text{if } n_i < n_j \\ 0.5 - (n_i^2 n_j + n_i n_j^2) & \text{if } n_i > n_j \end{cases} \tag{2}$$

where n_i is the population share of statehood proponents and n_j is the population share of statehood opponents.[1] Figure 3 plots an example of this measure. The figure shows the values of *Demographic polarization* and *Demographic polarization (directional)* as n_i changes for the case where $n_j = 1 - n_i$. When $n_i < n_j$, the two measures are identical. On the range where $n_i > n_j$, the directional measure of polarization continues to increase and increases at a faster rate as the difference in the group's population shares becomes more dramatic. I also code a directional measure

[1] There are no cases of these groups having equal population shares.

of cultural polarization:

$$\textit{Cultural polarization (directional)} = \begin{cases} (n_i^2 n_j + n_i n_j^2)(1 - s_{ij}^{\delta}) & \text{if } n_i < n_j \\ 0.5 - (n_i^2 n_j + n_i n_j^2)(1 - s_{ij}^{\delta}) & \text{if } n_i < n_j \end{cases} \tag{3}$$

where n_i is the population share of statehood proponents and n_j is the population share of statehood opponents. s_{ij} is defined as previously and $\delta = 0.5$.

Table 5 uses the directional measure of demographic polarization (Model 3) and cultural polarization (Model 4) in a reanalysis of the incidence of peaceful accommodation and violence. Both the linear and squared terms for relative INC representation are statistically significant in the equations for peaceful accommodation while only the squared term is statistically significant in the equations for violence. The implications of the estimated coefficients are better understood, however, by graphing the predicted relationships between relative INC representation, peaceful accommodation, and violence. Figure 4 plots the results from Model 3, which includes the directional measure of demographic polarization. The predicted relationships support the paper's argument even more sharply than the results in the main text. When relative INC representation is low—the pro-statehood interests were strongly favored in Congress—the predicted probability of peaceful accommodation is highest. At levels of relative INC representation clustered around one—parity between the interests on either side of the statehood debate—the predicted probability of violence is highest. At high levels of relative representation—the pro-statehood interests were strongly disfavored in Congress—both the predicted probability of violence and the predicted probability of peaceful accommodation are very low. Figure 5 displays very similar results based on the regression using the directional measure of cultural polarization as a control (Model 4).

The logistic regression analysis of accommodation in the main text controlled for polarization. Here, I check whether those results are robust to different measures of polarization. Table 6, Model 5 uses *Cultural polarization II* in the analysis of the overall incidence of accommodation as a function of violence and the interaction of violence and relative INC representation; Models 6 and

Table 5: Alternative measures of polarization: Reanalysis of peaceful accommodation and violence

	Model 2		Model 3		Model 4	
	Peaceful statehood	Violence	Peaceful statehood	Violence	Peaceful statehood	Violence
Ln relative INC representation	-4.4*	0.70*	-5.2*	-0.36	-5.0*	-0.52
	(2.1)	(0.28)	(2.3)	(0.60)	(2.4)	(0.66)
Ln relative INC representation sq.	-1.1*	-0.39*	-1.3*	-0.57*	-1.2*	-0.66*
	(0.44)	(0.11)	(0.50)	(0.21)	(0.50)	(0.21)
Cultural polarization II	7.5	0.58				
	(16)	(17)				
Demographic polarization (directional)			14	19		
			(8.7)	(12)		
Cultural polarization (directional)					28*	22*
					(16)	(12)
Ln enclave plurality group's INC rep.	-0.71	-1.4*	-1.0	-2.0*	-0.88	-1.8*
	(0.83)	(0.44)	(1.0)	(0.63)	(1.0)	(0.49)
Ln enclave plurality group's population	1.0	3.1*	1.2	3.5*	1.1	3.4*
	(0.97)	(0.60)	(1.2)	(0.69)	(1.1)	(0.60)
Agricultural labor share in enclave	-5.3*	-2.1	-6.5*	-2.6	-6.7*	-3.1
	(2.1)	(3.5)	(2.5)	(3.1)	(2.4)	(2.7)
Landless rate in enclave	11*	-7.2*	13*	-6.5*	12*	-6.6*
	(4.2)	(3.7)	(4.4)	(3.9)	(4.3)	(3.9)
Hindu share in enclave	2.5	-3.4*	4.0	-2.1	3.6	-2.5*
	(2.5)	(1.9)	(3.0)	(1.4)	(2.7)	(1.4)
Ln km to New Delhi	0.43	1.5*	0.62	2.1*	0.58*	1.8*
	(0.27)	(0.52)	(0.38)	(0.55)	(0.35)	(0.47)
Constant	-22*	-48*	-28*	-59*	-26*	-54*
	(12)	(11)	(16)	(14)	(15)	(11)
Observations	63		63		63	
Ln likelihood	-36		-34		-34	

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

Figure 4: Predicted probability of peaceful accommodation and violence as a function of relative INC representation of opponents and proponents of statehood (Based on Table 5, Model 3)

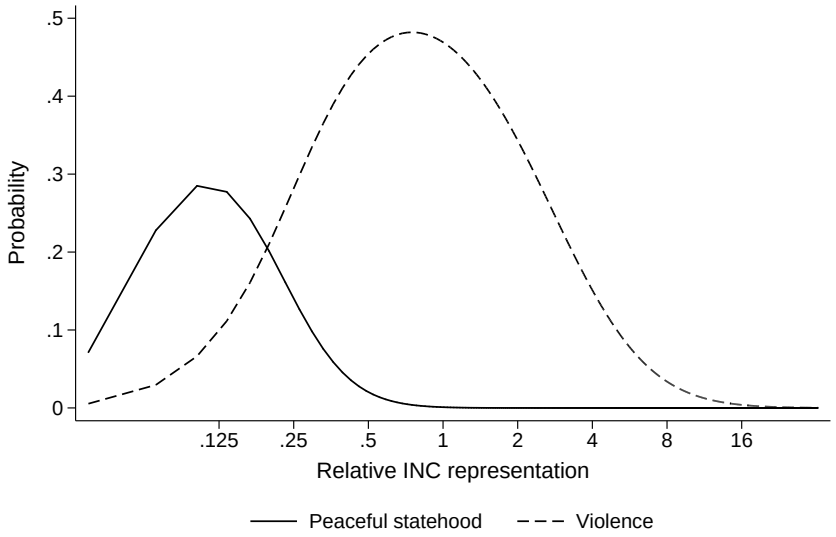
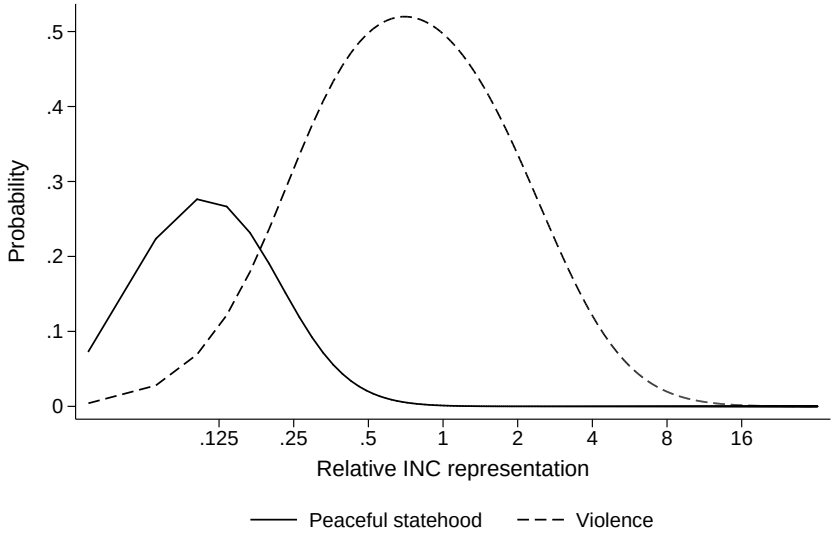


Figure 5: Predicted probability of peaceful accommodation and violence as a function of relative INC representation of opponents and proponents of statehood (Based on Table 5, Model 4)



7 use directional measures of polarization. In all three models, violence and the violence/relative representation interaction term have positive and statistically significant estimated coefficients, of similar magnitude to those in the main text.

5.2 Enclave-level polarization and fractionalization

All of the polarization measures introduced so far are calculated from the population shares of the enclave-level plurality language group and the largest group opposed to the enclave becoming a state. An alternative measurement strategy is to use the population shares of all the languages in each enclave to calculate an overall measure of polarization. Table 7 repeats the multinomial analysis including a measure of overall polarization for enclaves rather than concentrating on the two main opposed groups.

Model 8 incorporates enclave demographic polarization:

$$\textit{Enclave demographic polarization} = \Sigma_{i=1}^{m}\Sigma_{j\neq i} n_i^2 n_j \tag{4}$$

where n_i and n_j are the population shares of all language groups in the enclave, indexed 1 to m. Model 9 incorporates enclave cultural polarization:

$$\textit{Enclave cultural polarization} = \Sigma_{i=1}^{m}\Sigma_{j\neq i} n_i^2 n_j (1 - s_{ij}^{\delta}) \tag{5}$$

where s_{ij} is the number of common branches in a universal table of language genealogy divided by 15 and $\delta = 0.5$. Finally, I calculate ethnic fractionalization in each ethnic group's enclave:

$$\textit{Enclave fractionalization} = 1 - \Sigma_{i=1}^{m} n_i^2 \tag{6}$$

where n_i is an ethnic group's population share and m is the number of groups. This measure appears in Model 10.

Table 6: Alternative measures of polarization: Reanalysis of accommodation

	Model 5	Model 6	Model 7
	Statehood	Statehood	Statehood
Violence * Ln relative INC rep.	0.74*	0.97*	0.82*
	(0.32)	(0.31)	(0.30)
Violence	1.6*	1.7*	1.6*
	(0.84)	(0.90)	(0.86)
Ln relative INC representation	-0.032	0.44	0.28
	(0.18)	(0.49)	(0.45)
Cultural polarization II	0.80		
	(11)		
Demographic polarization (directional)		-9.7*	
		(5.3)	
Cultural polarization (directional)			-5.9
			(4.6)
Ln enclave plurality group's INC rep.	0.28	0.47	0.33
	(0.53)	(0.50)	(0.53)
Ln enclave plurality group's population	0.64	0.75	0.73
	(0.80)	(0.84)	(0.85)
Hindu share in enclave	1.9	0.59	1.6
	(1.5)	(1.9)	(1.6)
Ln km to New Delhi	1.1*	0.97*	1.1*
	(0.53)	(0.54)	(0.54)
Constant	-20	-18	-20
	(12)	(13)	(13)
Observations	63	63	63
Ln likelihood	-25	-24	-25

Standard errors in parentheses

Standard errors clustered by state

* $p < 0.10$

Table 7: Enclave-level polarization and fractionalization: Reanalysis of peaceful accommodation and violence

	Model 8		Model 9		Model 10	
	Peaceful statehood	Violence	Peaceful statehood	Violence	Peaceful statehood	Violence
Ln relative INC representation	-4.1*	0.80*	-4.4*	0.68*	-4.4*	0.98*
	(2.1)	(0.34)	(2.0)	(0.27)	(2.2)	(0.40)
Ln relative INC representation sq.	-1.1*	-0.41*	-1.1*	-0.38*	-1.2*	-0.44*
	(0.48)	(0.13)	(0.46)	(0.10)	(0.49)	(0.13)
Enclave demographic polarization	4.0	2.4				
	(3.3)	(3.2)				
Enclave cultural polarization			5.5	-1.2		
			(11)	(6.9)		
Enclave fractionalization					-3.1	-4.3
					(4.3)	(3.4)
Ln enclave plurality group's INC rep.	-0.50	-1.3*	-0.53	-1.5*	-0.85	-1.5*
	(1.3)	(0.50)	(1.1)	(0.47)	(0.98)	(0.51)
Ln enclave plurality group's population	1.4	3.4*	1.1	3.1*	1.1	3.1*
	(1.4)	(0.82)	(1.1)	(0.58)	(1.1)	(0.63)
Agricultural labor share in enclave	-4.8*	-0.93	-4.7*	-2.3	-5.7*	-2.8
	(2.3)	(3.9)	(2.1)	(3.4)	(2.3)	(3.4)
Landless rate in enclave	13*	-6.2	12*	-7.0*	13*	-6.1
	(4.7)	(4.3)	(4.3)	(3.6)	(4.2)	(4.5)
Hindu share in enclave	3.9	-3.7*	3.2	-3.4*	2.8	-3.3*
	(3.1)	(1.3)	(3.3)	(1.3)	(2.5)	(1.1)
Ln km to New Delhi	0.43	1.6*	0.39	1.5*	0.45	1.7*
	(0.29)	(0.49)	(0.31)	(0.50)	(0.29)	(0.51)
Constant	-29	-54*	-24*	-47*	-22	-46*
	(19)	(14)	(14)	(11)	(14)	(11)
Observations	63		63		63	
Ln likelihood	-35		-36		-35	

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

Table 8: Enclave-level polarization and fractionalization: Reanalysis of accommodation

	Model 11	Model 12	Model 13
	Statehood	Statehood	Statehood
Violence * Ln relative INC rep.	0.81*	0.76*	0.86*
	(0.45)	(0.37)	(0.31)
Violence	1.8*	1.7*	1.7*
	(1.1)	(0.96)	(0.80)
Ln relative INC representation	0.11	0.0056	-0.30
	(0.18)	(0.16)	(0.24)
Enclave demographic polarization	3.4		
	(2.5)		
Enclave cultural polarization		6.5	
		(7.2)	
Enclave fractionalization			3.4
			(4.1)
Ln enclave plurality group's INC rep.	0.83	0.74	0.24
	(0.84)	(0.89)	(0.54)
Ln enclave plurality group's population	0.60	0.55	0.83
	(0.75)	(0.76)	(0.90)
Hindu share in enclave	2.7*	2.2	1.7
	(1.5)	(1.5)	(1.7)
Ln km to New Delhi	1.2*	1.1*	0.99*
	(0.53)	(0.50)	(0.54)
Constant	-21*	-19*	-23
	(10)	(11)	(15)
Observations	63	63	63
Ln likelihood	-24	-24	-25

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

Adding these enclave measures of total demographic or cultural polarization or fractionalization to the analysis has no remarkable effects on the coefficients on the main variables. In Table 7, the terms for relative representation remain statistically significant across all equations and of similar magnitude to the coefficients reported in the main text.

I use the measures of enclave polarization and fractionalization in equations for accommodation in Table 8, Models 11–13. The coefficients on violence and the interaction term of relative representation and violence are positive and statistically significant in all three models.

5.3 Colonial-era grievances

Table 9 examines colonial treatment of the groups at odds over statehood. I collected new measures of language group incorporation in the colonial administrative apparatus. The first concerns colonial military service. After the 1857 Sepoy Mutiny, the British military began communal segregation of the military and narrowed recruitment to only a few ethnic groups (Cohen, 1971; Omissi, 1991). I use a 1919 British military review of recruitment (Hudson, 1919) to identify groups targeted for military service. Second, I code languages used by British civilian administration. Until the 1830s, Persian and English were the administrative languages of British-controlled India. After the 1830s, the British East India Company and, later, the British government theoretically encouraged the adoption of local vernaculars, along with English, as administrative languages (Mir, 2006); in practice, English remained dominant. In 1855, the East India Company librarian published a glossary of administrative terms in Indian languages that were sufficiently well-known to the British to be considered for local administration (Wilson, 1855). I use this glossary to code which languages were emphasized by the colonial service.

Based on my codings of colonial policy, I constructed measures comparing enclave plurality groups to their opponents in terms of colonial treatment. *Colonial military different* is a dummy variable coded as one if the language groups at odds were differentially treated by colonial military recruiters; *Colonial military disadvantaged* is coded as a one in the case that the enclave plurality

Table 9: Colonial-era grievances: Reanalysis of peaceful accommodation and violence

	Model 14		Model 15	
	Peaceful statehood	Violence	Peaceful statehood	Violence
Ln relative INC representation	-7.8*	0.66*	-6.0*	0.55*
	(4.1)	(0.25)	(2.7)	(0.20)
Ln relative INC representation sq.	-1.8*	-0.31*	-1.4*	-0.33*
	(0.87)	(0.095)	(0.54)	(0.086)
Colonial military different	-5.3*	-1.2		
	(2.0)	(0.98)		
Colonial admin. different	-0.76	0.56		
	(1.3)	(1.1)		
Colonial military disadvantaged			-4.1*	-1.1
			(1.7)	(0.90)
Colonial admin. disadvantaged			0.25	-0.86
			(1.1)	(0.97)
Demographic polarization	19*	28*	18*	20
	(11)	(15)	(11)	(15)
Ln enclave plurality group's INC rep.	-2.0*	-1.9*	-1.9*	-1.5*
	(0.84)	(0.58)	(0.76)	(0.57)
Ln enclave plurality group's population	1.9*	3.8*	1.8*	3.2*
	(1.00)	(0.71)	(0.90)	(0.57)
Agricultural labor share in enclave	-14*	-3.5	-13*	-3.4
	(4.3)	(5.1)	(4.1)	(4.7)
Landless rate in enclave	19*	-7.3	16*	-6.4
	(8.9)	(4.9)	(6.5)	(5.0)
Hindu share in enclave	9.7*	-2.4	7.2	-1.9
	(5.8)	(1.8)	(4.4)	(1.6)
Ln km to New Delhi	1.1*	2.8*	0.91*	2.3*
	(0.38)	(0.74)	(0.35)	(0.61)
Constant	-43*	-68*	-38*	-55*
	(14)	(15)	(13)	(13)
Observations	63		63	
Ln likelihood	-31		-33	

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

group was not recruited to the colonial military after 1857 while the opponents of statehood were recruited. *Colonial admin. different* notes whether the language groups at odds had experienced dissimilar colonial treatment with regard to the language of civil administration. *Colonial admin. disadvantaged* is coded as a one if the enclave plurality group's language was not used by the colonial administration while the opposed group's language was used.

In Table 9, the dummy variables for colonial treatment are incorporated in a multinomial analysis of violence and peaceful accommodation (Models 14 and 15). The coefficients on relative INC representation remain statistically significant and similar in magnitude to those in the main text.

5.4 State polarization, fractionalization, wealth, and inequality

Models 16–18 (Table 10) check for confounding effects of state demography, income, and inequality on the analysis of peaceful accommodation and violence. All three models include a count of the total number of statehood claims in the state where a group is located (*Total claims in state*), to capture the overall political pressure to split the state and the opportunity for contagion among violent movements. For Model 16, I calculate *State demographic polarization* using equation 4 (above) and the state population shares of all languages. Model 17 calculates *State cultural polarization* based on equation 5. Model 18 includes a measure of state-level ethnic fractionalization (*State fractionalization*), calculated based on equation 6.

All three models also control for state-level development—the percent of the workforce in agriculture (*Agricultural labor share in state*)—and for state-level inequality, which is the landless share of the agricultural workforce (*Landless rate in state*).

In all three models in Table 10, the coefficients on relative INC representation, peaceful accommodation, and violence are statistically significant. The estimated coefficients on relative representation in the equation for peaceful accommodation are larger than those in the main text.

Table 10: State polarization, fractionalization, wealth, and inequality: Reanalysis of peaceful accommodation and violence

	Model 16		Model 17		Model 18	
	Peaceful statehood	Violence	Peaceful statehood	Violence	Peaceful statehood	Violence
Ln relative INC representation	-13*	1.2*	-13*	1.2*	-13*	1.2*
	(4.6)	(0.34)	(4.6)	(0.34)	(4.5)	(0.35)
Ln relative INC representation sq.	-3.1*	-0.35*	-3.1*	-0.35*	-3.1*	-0.37*
	(1.1)	(0.13)	(1.1)	(0.13)	(1.1)	(0.12)
Total claims in state	-1.1*	-0.16	-1.1*	-0.16	-1.1*	-0.19
	(0.37)	(0.20)	(0.37)	(0.20)	(0.37)	(0.20)
State demographic polarization	-1.7	-7.8				
	(3.9)	(6.1)				
State cultural polarization			-1.7	-7.8		
			(3.9)	(6.1)		
State fractionalization					1.8	8.3
					(4.2)	(6.0)
Agricultural labor share in state	-1.4	-8.5	-1.4	-8.5	-1.3	-8.2
	(8.3)	(8.9)	(8.3)	(8.9)	(8.3)	(8.6)
Landless rate in state	24	-4.9	24	-4.9	24	-4.3
	(29)	(12)	(29)	(12)	(29)	(12)
Ln enclave plurality group's INC rep.	-2.5*	-2.1*	-2.5*	-2.1*	-2.5*	-2.1*
	(1.0)	(0.78)	(1.0)	(0.78)	(0.99)	(0.77)
Ln enclave plurality group's population	2.2*	4.0*	2.2*	4.0*	2.2*	4.1*
	(0.82)	(0.83)	(0.82)	(0.83)	(0.79)	(0.82)
Agricultural labor share in enclave	-8.6	-0.32	-8.6	-0.32	-8.6	-0.55
	(5.5)	(4.6)	(5.5)	(4.6)	(5.4)	(4.7)
Landless rate in enclave	-0.52	-8.3	-0.52	-8.3	-0.62	-8.5
	(20)	(12)	(20)	(12)	(20)	(12)
Hindu share in enclave	-3.3	-3.2	-3.3	-3.2	-3.4	-3.3
	(3.8)	(2.4)	(3.8)	(2.4)	(3.8)	(2.4)
Ln km to New Delhi	0.99*	1.9*	0.99*	1.9*	0.99*	2.0*
	(0.48)	(0.58)	(0.48)	(0.58)	(0.48)	(0.57)
Constant	-38*	-54*	-38*	-54*	-39*	-62*
	(13)	(12)	(13)	(12)	(12)	(11)
Observations	63		63		63	
Ln likelihood	-28		-28		-28	

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

5.5 State versus enclave inequality

I also compare agricultural dependence in each group's area to that of their state (Table 11), constructing a measure of horizontal inequality suggested by Cederman et al. (2011). If a is the enclave's rate of agricultural employment and A is the state's:

$$\textit{Inequality (State vs. enclave)} = [ln(a/A)]^2 \tag{7}$$

Enclaves with high or low rates of agricultural employment relative to the state will have positive scores for inequality using this measure. This measure is used in Model 19, a multinomial logistic regression of peaceful accommodation and violence, in Table 11.

An alternate measure of horizontal inequality distinguishes enclaves that are better or worse off than the state average:

$$\textit{Relative wealth (State vs. enclave)} = \begin{cases} A/a & \text{if } a < A \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

$$\textit{Relative deprivation (State vs. enclave)} = \begin{cases} a/A & \text{if } a > A \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

For an enclave with higher agricultural employment than the state as a whole, *Relative deprivation* is a positive number greater than one and *Relative wealth* is zero. For an enclave with lower agricultural employment than the state as a whole, *Relative deprivation* is zero and *Relative wealth* is a positive number greater than one. These variables are used in Model 20 in Table 11.

In both Models 19 and 20, the linear and squared terms for relative INC representation remain statistically significant in the equations for peaceful accommodation and violence. The estimated coefficients on relative representation in the equation for peaceful accommodation are larger than those in the main text.

Table 11: State versus enclave inequality: Reanalysis of peaceful accommodation and violence

	Model 19		Model 20	
	Peaceful statehood	Violence	Peaceful statehood	Violence
Ln relative INC representation	-13*	1.3*	-14*	1.7*
	(5.1)	(0.42)	(6.4)	(0.41)
Ln relative INC representation sq.	-3.2*	-0.39*	-3.4*	-0.52*
	(1.1)	(0.14)	(1.4)	(0.19)
Inequality (State vs. enclave)	-83*	-16		
	(47)	(37)		
Relative deprivation (State vs. enclave)			-12	-7.0
			(11)	(9.1)
Relative wealth (State vs. enclave)			-13	-10
			(11)	(8.9)
Agricultural labor share in state	-5.8	-11	-2.3	7.2
	(8.2)	(11)	(10)	(14)
Landless rate in state	17	-6.9	25	-9.2
	(30)	(12)	(32)	(11)
State cultural polarization	3.5	-6.0	4.1	-6.4
	(3.0)	(5.7)	(3.0)	(5.5)
Ln enclave plurality group's INC rep.	-2.4*	-1.9*	-2.4*	-2.2*
	(0.96)	(0.66)	(1.2)	(0.94)
Ln enclave plurality group's population	1.8*	3.9*	1.8*	4.9*
	(0.68)	(0.72)	(0.80)	(1.2)
Agricultural labor share in enclave	-7.4	2.2	-10	-17
	(7.1)	(7.2)	(9.8)	(15)
Landless rate in enclave	4.1	-6.0	0.56	-9.2
	(22)	(12)	(21)	(12)
Hindu share in enclave	3.7	-3.0	2.4	-7.3*
	(5.5)	(2.4)	(5.6)	(2.8)
Ln km to New Delhi	0.90*	1.9*	0.74	2.6*
	(0.47)	(0.53)	(0.47)	(0.63)
Constant	-38*	-54*	-26*	-58*
	(10)	(11)	(9.3)	(17)
Observations	63		63	
Ln likelihood	-30		-28	

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

5.6 State politics

Recent scholarship shows a strong relationship between India's state-level politics and ethnic mobilization or violence. However, in the 1950s, state politics was not nearly as competitive as national politics or as state politics in later decades. The Congress party controlled all but one state government—PEPSU was the exception—and only one state had a coalition government, the Congress-led government in Travancore-Cochin. Nonetheless, this section investigates the robustness of earlier findings to the addition of state-level political controls. Enclaves in type C states that did not have elected state governments are not included in this analysis.

Wilkinson (2004) shows two effects of political competition on religious riots in India. At the municipality level, electoral competition exacerbates violence. At the state level, a multiparty system reduces violence. I measure enclave-level electoral competitiveness, calculating the average margin of victory in all 1951 races for parliament (*Average winner's vote share in enclave*). I measure the effective number of political parties in each state's Vidhan Sabha (state assembly), based on seat shares.[2] This measure is called *Effective parties in Vidhan Sabha*. These two variables are included in multinomial analysis of peaceful accommodation and violence in Model 21, Table 12.

Chandra (2004) argues that ethnic parties emerge when the dominant political party or parties cannot absorb ambitious elites from minority groups. I construct two measures of the Congress' ability to absorb politically ambitious newcomers. The first is the strength of the state Congress party, measured as its seat share in the Vidhan Sabha in 1951 (*INC seat share in Vidhan Sabha*). The second codes whether a language group had a designated Pradesh (province) Congress Committee after the reorganization of Congress in 1921 (*PCC since 1921*) (see Krishna, 1966, 415–16). The 1921 reorganization of Congress converted the PCCs into language-based units, although each PCC was confined to a particular state. For example, the Maharashtra PCC represented Marathi speakers in Bombay; the Central Provinces Marathi PCC represented Marathi speakers in the Central Provinces. *PCC since 1921* is coded as a one if an enclave had a PCC in 1921. So, this variable

[2]Election data based on ECI (1951) and ECI (2012).

Table 12: State politics: Reanalysis of peaceful accommodation and violence

	Model 21		Model 22	
	Peaceful statehood	Violence	Peaceful statehood	Violence
Ln relative INC representation	-8.8*	0.77*	-9.1*	1.2*
	(5.4)	(0.39)	(3.5)	(0.44)
Ln relative INC representation sq.	-2.0*	-0.46*	-2.0*	-0.52*
	(1.2)	(0.23)	(0.66)	(0.18)
Average winner's vote share in enclave	2.4	-11		
	(7.5)	(11)		
Effective parties in Vidhan Sabha	1.2*	-0.22		
	(0.70)	(0.31)		
INC seat share in Vidhan Sabha			-18*	-3.6
			(5.6)	(3.2)
PCC since 1921			0.85	3.8*
			(1.0)	(1.5)
Demographic polarization	5.0	6.6	-3.3	9.5
	(14)	(14)	(13)	(13)
Ln enclave plurality group's INC rep.	0.22	-1.5*	1.4	-2.0*
	(1.1)	(0.67)	(1.2)	(0.75)
Ln enclave plurality group's population	0.042	3.2*	-0.62	3.9*
	(1.1)	(0.87)	(1.2)	(0.98)
Agricultural labor share in enclave	-2.3	-0.69	-3.6	3.5
	(2.1)	(5.9)	(2.6)	(5.0)
Landless rate in enclave	11*	0.048	16*	-4.7
	(3.7)	(5.6)	(5.4)	(8.8)
Hindu share in enclave	4.0	-3.2*	9.3	-3.6*
	(3.6)	(1.7)	(5.8)	(1.5)
Ln km to New Delhi	0.18	1.8*	0.22	2.2*
	(0.44)	(0.53)	(0.44)	(0.70)
Constant	-19	-47*	0.047	-66*
	(19)	(16)	(19)	(19)
Observations	59		59	
Ln likelihood	-30		-25	

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

is coded as a one for Marathis in Bombay and Madhya Pradesh (née Central Provinces) but not for Marathis in Hyderabad, where there was no pre-independence PCC. Model 22 incorporates *INC seat share in Vidhan Sabha* and *PCC since 1921*.

The main variables remain statistically significant in Models 21 and 22. As in other models using state-level controls, the coefficients on relative representation in the equation for peaceful accommodation are larger than the estimates in the main text.

5.7 Cultural distance and separatism

State reorganization violence may also have been related to groups' separatist tendencies or antipathy toward Indian nationalism. Such an antipathy could lead to both violence and to low vote shares for the Congress party. One factor that mitigates this problem is that the analysis excludes peripheral areas that were not fully under New Delhi's control in the 1950s (i.e., Jammu and Kashmir, the Northeast Frontier Administration (NEFA), and the Andaman and Nicobar Islands). These were the clearest separatist threats of the time. Also, all regressions control for distance from New Delhi and religious composition, which should capture peripherality.

Table 13 adds additional controls for possible separatism. The first concerns religious difference. The regressions in the main text already distinguish religious minority areas from Hindu areas. However, Hinduism is very internally diverse. The dominant variety of Hinduism in Indian nationalist rhetoric is Brahminical Hinduism, which is based in north India. Jha (2012) argues that proximity to the Ganges river indicates how Brahminical the religious and caste systems in different regions of India are; this relationship is due to beliefs regarding sacred geography. I code *Ln km to Varanasi*, the most important religious site on the Ganges, as a measure of dissimilarity from Brahminical religious traditions.

A dummy variable for a *Land or sea border* captures the inherent feasibility of independence.

I also include a voting-based measure of support for regionalism/separatism. The first is the Lok Sabha vote share in the area won by political parties with regionalist agendas (*Regionalist*

Table 13: Separatist tendencies: Reanalysis of peaceful accommodation, violence, and accommodation

	Model 23		Model 24		Model 25	Model 26
	Peaceful statehood	Violence	Peaceful statehood	Violence	Statehood	Statehood
Ln relative INC representation	-3.9	0.51*	-4.7*	0.43*	-0.31	-0.13
	(2.5)	(0.25)	(2.7)	(0.24)	(0.62)	(0.39)
Ln relative INC representation sq.	-0.82*	-0.40*	-0.99*	-0.35*		
	(0.48)	(0.10)	(0.56)	(0.10)		
Violence * Ln relative INC rep.					2.6*	1.7*
					(1.0)	(0.40)
Violence					6.5*	2.9*
					(2.9)	(1.0)
Ln km to Varanasi	1.8*	0.17	1.2	-0.081	3.8*	2.4*
	(0.91)	(0.80)	(0.99)	(0.82)	(1.3)	(1.1)
Land or sea border	-2.7	0.56	-2.9	0.13	-3.9*	-2.8*
	(2.2)	(0.98)	(2.1)	(0.89)	(1.8)	(1.2)
Regionalist party vote share in enclave	-12*	6.5			-33*	
	(7.4)	(4.6)			(9.7)	
Hindu right vote share in enclave			-5.7	-16		-11
			(6.7)	(10)		(10)
Demographic polarization	10	15	13	21*	-13*	-16*
	(9.3)	(12)	(9.4)	(12)	(5.5)	(8.4)
Ln enclave plurality group's INC rep.	-0.81	-1.8*	-1.0	-1.8*	0.44	0.50
	(1.4)	(0.64)	(1.1)	(0.58)	(0.75)	(0.76)
Ln enclave plurality group's population	1.5	3.3*	1.6	3.3*	1.6	1.2
	(1.5)	(0.61)	(1.3)	(0.56)	(1.1)	(1.1)
Agricultural labor share in enclave	-5.7*	-4.6	-8.0*	-2.6		
	(3.3)	(4.4)	(3.2)	(3.9)		
Landless rate in enclave	5.4	-5.3	4.2	-11*		
	(9.1)	(5.3)	(8.7)	(6.4)		
Hindu share in enclave	-3.3	0.99	1.9	-0.100	-7.6*	1.5
	(5.7)	(2.3)	(3.1)	(1.8)	(3.8)	(2.5)
Ln km to New Delhi	0.38	2.0*	0.43	1.8*	-0.096	-0.17
	(0.46)	(0.49)	(0.47)	(0.50)	(0.29)	(0.37)
Constant	-34*	-59*	-35*	-55*	-41*	-33
	(20)	(13)	(21)	(11)	(17)	(21)
Observations	63		63		63	63
Ln likelihood	-32		-31		-16	-18

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

party vote share in enclave).[3] Second, I code the vote share in each enclave for Hindu nationalist parties (*Hindu right vote share in enclave*).[4] In addition to being the most aggressively nationalist political parties, the Hindu parties opposed the linguistic reorganization of states on national security grounds. High Hindu nationalist vote shares signify the popularity of one type of Indian nationalism.

Models 23 and 24 use these measures of cultural distance, opportunity for separatism, and support for regionalists or right-wing nationalists in multinomial models of peaceful accommodation and violence. The coefficients on relative INC representation have the same magnitude and sign as the results seen in the main text. In Model 23, the linear term for relative representation is not statistically significant in the equation for peaceful accommodation (p=0.12). The statistical significance for all other relative representation terms remains.

When language movements were perceived by New Delhi as separatist or potentially separatist, these movements were less likely to be accommodated (Brass, 1974). Therefore, I reestimate the role of violence as a determinant of accommodation controlling for the additional separatism-related variables in Table 13, Models 25 and 26. Violence and the interaction term between violence and relative INC representation are both positively signed and statistically significant.

5.8 Differential media coverage

Finally, there may be biases in the data collected for this paper due to the reliance on an English-language newspaper, the *Times of India*, published in Bombay City. Events occurring closer to Bombay were both more likely to be known to the newspaper's staff and more likely to be reported because of greater reader interest. To some extent, this problem is mitigated by the exclusion of

[3]The regionalist parties are the All Manipur National Union, Chota Nagpur Santhal Parganas Janta Party, Cochin Party, Punjab Depressed Class League, Hill People Party, Hyderabad State Praja Party, Jharkhand Party, Khasi-Jaintia Durbar, Kuki National Association, Kerala Socialist Party, Madras State Muslim League Party, Uttar Pradesh Revolutionary Socialist Party, Shiromani Akali Dal, Saurashtra Khedut Sangh, Tamil Nad Toilers Party, Tamil Nad Congress, Tribal Sangha, Travancore Tamil Nad Congress Party, Uttar Pradesh Praja Party.

[4]The All India Bhartiya Jan Sangh, the Akhil Bharatiya Hindu Mahasabha, and the Akhil Bharatiya Ram Rajya Parishad.

NEFA, Jammu and Kashmir, and the Andaman and Nicobar Islands from the data. These were the areas least open to journalists due to restrictions on the press and the lack of government contact with the most remote areas.

Table 14 introduces controls for biases in *Times of India* coverage. Model 27, for peaceful accommodation and violence, controls for each enclave's *Ln km to Bombay* and the count of daily newspapers, both vernacular and English-language, in each enclave. I code *Vernacular daily newspapers in enclave* and *English daily newspapers in enclave* based the 1957 *Annual Report of the Registrar of Newspapers for India* (MIB, 1957). The street address for each paper is listed in the report, allowing me to match the papers to enclaves as they were prior to reorganization.[5]

In Model 27 the results on the main explanatory variables are similar to those in the main text and remain statistically significant.

I also re-estimate the relationship between violence and accommodation using controls for media coverage. Differential media coverage has the potential to introduce biases in my measurement of violence, possibly skewing the estimates in the main text of the relationship between violence and accommodation. However, violence and the interaction term between violence and relative INC representation have positive and statistically significant coefficients in a model including controls for print media (Table 14, Model 28).

[5]The dates of each newspaper's founding are not included in the report, so it is possible that some of these papers began publication after state reorganization was completed in 1956; conversely, some papers that were publishing before 1957 may have closed before the report was published.

Table 14: Print media: Reanalysis of peaceful accommodation, violence, and accommodation

	Model 27		Model 28
	Peaceful statehood	Violence	Statehood
Ln relative INC representation	-4.9*	0.57*	0.056
	(2.6)	(0.34)	(0.20)
Ln relative INC representation sq.	-1.2*	-0.29*	
	(0.57)	(0.15)	
Violence * Ln relative INC rep.			1.1*
			(0.54)
Violence			1.9*
			(0.86)
Ln km to Bombay	-0.88	-0.29	-1.6*
	(1.1)	(1.0)	(0.75)
Vernacular daily newspapers in enclave	0.0019	0.16	0.060
	(0.099)	(0.13)	(0.093)
English daily newspapers in enclave	0.15	-0.024	0.52*
	(0.18)	(0.38)	(0.26)
Demographic polarization	15*	19	-11
	(8.4)	(17)	(11)
Ln enclave plurality group's INC rep.	-0.98	-2.3*	0.11
	(0.81)	(0.92)	(0.41)
Ln enclave plurality group's population	1.1	3.0*	0.38
	(1.1)	(0.90)	(0.74)
Agricultural labor share in enclave	-6.2*	-1.4	
	(2.8)	(4.6)	
Landless rate in enclave	8.7	-7.3	
	(5.7)	(9.5)	
Hindu share in enclave	2.8	-1.5	0.077
	(3.0)	(1.7)	(2.5)
Ln km to New Delhi	0.77*	2.3*	1.3*
	(0.38)	(0.87)	(0.73)
Constant	-19	-54*	-4.3
	(13)	(13)	(9.5)
Observations	63		63
Ln likelihood	-33		-20

Standard errors in parentheses
Standard errors clustered by state
* $p < 0.10$

References

Brass, Paul R. 1974. *Language, Religion and Politics in North India*. New York: Cambridge University Press.

Cederman, Lars-Erik, Nils B. Weidmann, and Kristian S. Gleditsch. 2011. "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison." American Political Science Review 105 (3): 478–495.

Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed: Patronage and Head Counts in India*. New York: Cambridge University Press.

Cohen, Stephen P. 1971. *The Indian Army: Its Contribution to the Development of a Nation*. Berkeley, CA: University of California Press.

ECI. 1951. *Delimitation of Parliamentary and Assembly Constituencies Order, 1951*. New Delhi: Election Commission of India.

ECI. 2012. *General Election Results and Statistics*. New Delhi: Election Commission of India.

Esteban, Joan, Laura Mayoral, and Debraj Ray. 2012. "Ethnicity and Conflict: Theory and Facts." Science 336 (6083): 858–865.

Hudson, H. 1919. *Recruiting in India Before and During the War of 1914–18*. Delhi: Army Headquarters, India.

Jha, Saumitra. 2012. "Trade, Institutions and Ethnic Tolerance: Evidence from South Asia." Stanford GSB Working Paper 2004.

Krishna, Gopal. 1966. "The Development of the Indian National Congress as a Mass Organization, 1918–1923." The Journal of Asian Studies 25 (3): 413–430.

MIB. 1957. *Annual Report of the Registrar of Newspapers for India*. New Delhi: Ministry of Information and Broadcasting, Government of India.

Mir, Farina. 2006. "Imperial Policy, Provincial Practices: Colonial Language Policy in Nineteenth-Century India." The Indian Economic and Social History Review 43 (4): 395–427.

Omissi, David. 1991. "'Martial Races:' Ethnicity and Security in Colonial India, 1858–1939." War and Society 9 (1): 1–27.

Wilkinson, Steven. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. New York: Cambridge University Press.

Wilson, H.H. 1855. *Glossary of Judicial and Revenue Terms and of Useful Words Occurring in Official Documents Relating to the Administration of the Government of British India*. London: W. H. Allen and Co.