Part

Online Appendix

Table of Contents

A More Contextual Discussion of ST, Forests, Resistance

A.1 The British Colonial Period and Scheduled Tribes

The ST identity category was first codified, with corresponding separate administrative areas specified, during the British Colonial period. Scholars have identified these 'tribal' groups by (a) their descent from particular lineages (Sundar 2009), (b) pre-colonial systems of administration, and/or (c) well-defined land arrangements and rights (Gupta 2011). Despite regular mention of these factors, scholars agree that there has been little clear definition or criteria as to what, precisely, constitutes an Indian 'tribe' (Béteille 1974; Skaria 1997; Corbridge, Jewitt, and Kumar 2004; Corbridge 2002; Galanter 1984).

Encountering what are now labeled ST communities, British administrators defined and enumerated those they viewed as 'tribal' populations. British authorities first provided a list of 'Aboriginal Tribes' and 'Semi-Hinduised Aboriginal Tribes' in the Census of 1872 (Corbridge 2002, p. 64) and introduced special institutions based on this census with the Scheduled Districts Act of 1874. These communities were not distributed randomly, but geographically concentrated in areas distant from urban areas that were heavily forested and hilly.

A.2 The Stability of Scheduled Areas since Independence

The geographic boundaries of areas the Scheduled Areas have changed relatively little over time. Per the Constitution, the President of India has the right to Schedule or De-schedule Areas, in consultation with State Governors. In 1962, the Dhebar Commission proposed that an area should be eligible to become a Scheduled Area according to the following four (vague) criteria: (a) Preponderance of tribals in the population, (b) Compact and reasonable size, (c) Under-developed nature of the area, and (d) Marked disparity in economic standards of the people (Dhebar 1962). In practice there has been no exact formula for updating or adjusting notification of Scheduled Areas, and these Areas have remained remarkably stable since the Dhebar Commission.[26]

A.3 Forest Policy and ST Resistance

ST have been disproportionately linked to forests and forest policy. The British set up an extractive institution of forestry management beginning in 1864 with the establishment of the Imperial Forest Department. The British forest Acts in 1865 and 1878, the latter of which, reproduced verbatim in 81 of the 84 sections of the Indian Forest Act of 1927 (Guha 2000), continues to shape forestry policy today. These acts consolidated exclusive state control over forests to meet the economic demand for timber – particularly driven by railroads and shipyards, in so doing alienating local communities (Sundar 2007). The colonial model of extraction continued after Independence, but the Indian state shifts its justification to commercial objectives and conservation imperatives (Gadgil and Guha 1992).

The British Acts set a precedent that was reproduced under the early Indian state, which in the Indian Forest Act of 1927 effectively reproduced, nearly in full, the earlier British Acts. This reinforced the Forest Department, as an institution working in opposition to, rather than alongside or in support of, ST and other rural forest communities (Guha 2000; Patnaik 2007).

Despite official recognition of Scheduled Tribes in Article 366 of the Indian Constitution, ST rights worsened post-Independence as large tracts of land were declared "forest" by land owners (*zamindari*), heads of princely states, and other private owners, through blanket government notifications (Patnaik 2007, p. 5). With the claim that ST and other forest dwellers had destroyed forest resources, and that those resources needed protection, the Wildlife Protection Act 1972 and the Forest Conservation Act 1980, brought these areas under the purview of the Forest Department.

ST have a history of conflict against the state, first against the British absorption of these communities right to use their lands and forests, and later against the Indian state and in particular the Forest Department. In waves, initially under British Rule, and with additional rounds of forest reservation by

[26]Gulzar, Haas, and Pasquale (2020) Appendix A reports additional details on the institutional background of Scheduled Areas as well as additional information on Scheduled Tribes.

the Indian state, implemented by the Forest Department, ST have been forcibly evicted from forests and lost their rights to collect for sustenance, as well as to sell, non-timber forest produce (Vasan 2009, p. 127; Shah 2013, pp. 431, 436).

A.4 Context in Jharkhand

The case of Jharkhand illuminates how ST contested incursions by the British and subsequent Indian government into the lands and forests among which they lived. From 1895-1900, a series of tribal revolts, the most famous of which was led by Birsa Munda and known as the great tumult (*Ulgulan*) protested the loss of community ownership rights to forests *khuntkatti*. The Birsa Munda revolt was not an isolate event but representative of many so-called 'tribal rebellions' on issues of land use and land alienation, focused in particular on restricting ST access to, and use of, forest products.[27] In response, the British acknowledged these initial ownership and rights by means of enacting the Chotanagpur Tenancy Act in 1908 as well as the Santhal Parganas Tenancy Act of 1949. Collectively, the land covered by these two tenancy acts covers exactly the districts that were later separated from the state of Bihar to form the new state of Jharkhand in 2000 (Shah 2010; Sundar 2009). These Acts prohibit transfer of land to non-tribals and aims to protect community ownership and management rights of forest communities in ST *khuntkatti* areas.

In Jharkhand, scholars studying ST in the post-Colonial and contemporary periods continue to see

[27]These include the Chipko movement (in what is now the state of Uttrakhand), where early uprisings against the British in the 1930s and 1940s were reborn with the forest conservation movement from 1973 to 1981, which inspired future environmental movements around the world (Guha 2000). For example, in the region of Bastar (within today's state of Chhattisgarh), initial rebellions in 1876 repeated in 1910 (Sundar 2007; Verghese 2016); in what is now Jharkhand, the Santhal Rebellion occurred in 1855. More recently, Kond communities organized protests in Kashipur block and Gandhamardan Hills regions (today part of Odisha), against the corporation Vedanta and their aluminum (bauxite) mines (Padel and Das 2010).

patterns of resistance. Vasan writes, "Forests, which are the lifeline of Adivasi livelihood, culture, and society...also been the primary target of Adivasi protest" (Vasan 2009, p. 113). Shah describes how non-state armed groups worked on behalf of ST, against state actors to gain ST support by: "bombing the state forest rest houses, burning forest jeeps, and chasing out the officers, coupled with Maoist replacement of outside contractors with locals, and raising the wages of forest product collection, have been extremely influential in gaining Adivasi support" (Shah 2010, p. 346).

In Jharkhand, the Government of India has not only restricted access to and alienated land from ST, but also encouraged industrial operation and mining operations in ST-dominated areas. Scholars have estimated that from 1951-1991, between 963,000 and 1.5m individuals have been displaced, approximately 28-43% of whom are ST. This displacement has been driven by more than 500,000 acres alienated due to mining projects and 500,000 acres were converted to national parks where individuals could no longer access their customary rights (Ekka and Asif 2000; Sharan 2009). As with respect to land and forest rights, ST and other communities have resisted these projects with protests against coal mines, dams, military facilities (Sundar 2009, p. 24).

B Measurement Error Issues

Since our outcome variable is a remote-sensing based prediction of forest cover, it is worth thinking through consequences of measurement error and data issues on our estimates. Jain (2020) discusses potential pitfalls in the use of remote-sensing data, and in particular focuses on the consequences of measurement error in raw satellite output and the importance of validation analyses. Errors in remote-sensing measures are likely 'non-classical' (that is, not mean-zero normally distributed) in nature. This may be caused by sensor characteristics, atmospheric conditions, cloud cover, and so on, since satellites may produce systematically worse measures of forest cover in specific places (e.g. cloudy or dusty locations). However, since our empirical strategy is longitudinal, while all systematic noise in the data is cross-sectional, these sources of error are partialled out by pixel and village fixed-effects. To see this, consider a location that is consistently poorly measured due to cloud cover, and another which is measured with minimal error. Pixel fixed-effects estimate a separate intercept for each of these pixels, where the intercept for the first pixel is artificially low due to measurement error, while it is accurate for the latter. However, as long as *changes* in forest cover are appropriately captured by satellite measures, biases in the estimation of the fixed-effects [28] do not affect the causal estimates, which depend on *within-unit* variation in forest cover over time. Hence, changes are likely accurately measured even if levels are not, and as such this is unlikely to be a major source of bias.

An alternate source of inconsistency is that older satellites get sunset, while newer, better ones come on line, thereby resulting in inconsistency in measurements over time. These are typically addressed in harmonization in pre-processing, as by Song et al. (2018) and Hansen et al. (2013), by fitting a polynomial trends and intercept shifts when new satellites come online. Additionally, since we estimate specifications with year and state-year fixed effects, we adjust for this source of error by letting each year have a separate intercept, which absorbs idiosyncrasies in measurement that are common across all pixels for a given year. The only way this could bias our estimates is if different locations were measured by different satellites within the same time period, which never happens.[29]

[28] In fact, it is well known that the estimates of fixed effects are inconsistent many standard settings thanks to the incidental parameters problem (Lancaster 2000).

[29] State-fixed effects even account for this remote possibility by permitting each state to have its own intercept.

C Summary Statistics

C.1 VCF Data

Table A1: Summary Statistics for primary analysis sample (VCF Data) - full sample

Variable	NotNA	Mean	Sd	Min	Pctile[25]	Median	Pctile[75]	Max
Forest cover index (0-100)	2204152	5.0873	8.3424	0	0	2	6	88
Non-forest green index (0-100)	2204152	68.732	21.297	0	64	77	83	99
Non-green index (0-100)	2204152	25.887	23.651	0	10	17	33	99
Scheduled Status	2204152	0.1752	0.38014	0	0	0	0	1
Forest Cover in 1990 (Ex-Ante)	2204152	5.4625	9.4535	0	0	2	6	85

Table A2: Summary Statistics for primary analysis sample (VCF Data) - above median forest cover in 1990

Variable	NotNA	Mean	Sd	Min	Pctile[25]	Median	Pctile[75]	Max
Forest cover index (0-100)	1048662	9.5032	10.319	0	3	6	13	88
Non-forest green index (0-100)	1048662	79.212	7.8119	4	77	81	84	98
Non-green index (0-100)	1048662	11.285	7.2883	0	6	10	14	92
Scheduled Status	1048662	0.2634	0.44048	0	0	0	1	1
Forest Cover in 1990 (Ex-Ante)	1048662	10.75	11.558	3	3	6	14	85

C.2 GFC Data

Table A3: Summary Statistics (GFC Data) - full sample

Variable	NotNA	Mean	Sd	Min	Pctile[25]	Median	Pctile[75]	Max
Deforested Area (Hectares)	4877470	0.045483	1.0379	0	0	0	0	351.63
Scheduled Status	4877470	0.18688	0.38982	0	0	0	0	1
Ex-ante forest cover in 2000 (ex-ante)	4877470	3.2006	9.2771	0	0.0069353	0.074147	0.69017	89.098

Table A4: Summary Statistics (GFC Data) - above 2 percent forest cover in 2000

Variable	NotNA	Mean	Sd	Min	Pctile[25]	Median	Pctile[75]	Max
Deforested Area (Hectares)	897192	0.22	2.4	0	0	0	0	352
Scheduled Status	897192	0.3	0.46	0	0	0	1	1
Ex-ante forest cover in 2000 (ex-ante)	897192	17	16	2	4.5	10	25	89

D Additional Results

D.1 Alternative Estimators

Our preferred specification includes cell/village fixed-effects, state-year fixed effects, and cell level time trends. While this is a very flexible specification, one might still be concerned about the functional form assumptions of linearity. In this section, we describe a framework for casting various disparate estimation strategies into a common framework, due to Doudchenko and Imbens (2016) and Athey et al. (2021).

We first write the outcome matrix as a $N \times T$ matrix, where the ith corresponds to the time series of outcomes for unit i

$$\mathbf{Y}^{\text{obs}} = \begin{bmatrix} Y_{1,\,1}^{\mathcal{T},\texttt{pre}} & \dots & Y_{1,\,\text{T}}^{\mathcal{T},\texttt{post}} \\ \vdots & \ddots & \vdots \\ Y_{\text{N},\,1}^{\mathcal{C},\texttt{pre}} & \dots & Y_{\text{N},\,\text{T}}^{\mathcal{C},\texttt{post}} \end{bmatrix}$$

with rows corresponding with units $i = 1, \dots, N$ columns corresponding with time periods $t = 1, \dots, T$, observations for ever-treated and never-treated groups before and after treatment indexed $\mathcal{T}, \texttt{pre}$, $\mathcal{T}, \texttt{post}$, $\mathcal{C}, \texttt{pre}$, and $\mathcal{C}, \texttt{post}$ respectively. The outcome matrix is generated from the underlying potential outcome matrices $\mathbf{Y}(0), \mathbf{Y}(1)$ as $\mathbf{Y}^{\text{obs}} = (1 - \mathbf{W})\mathbf{Y}(0) + \mathbf{W}\mathbf{Y}(1)$, where $\mathbf{W}$ is a $N \times T$ treatment indicator matrix. We seek to estimate the average treatment effect on the treated (ATT), which is $\mathbb{E}\left(\mathbf{Y}(1) - \mathbf{Y}(0) | i \in \mathcal{T}\right)$. Since potential outcomes $\mathbf{Y}(0)$ are missing for treated units after the treatment is implemented, we need to impute them. The two-way fixed effects model imposes a linear additive form on the untreated potential outcome $Y(0) = \delta_i + \gamma_t$, and assumes an additive treatment effect, but alternative imputation methods are possible.

Depending on the relative dimensions of the matrix $\mathbf{Y}^{\text{obs}}$, the outcome matrix may be a thin matrix ($T << N$), a fat matrix ($T >> N$), and approximately square ($N \approx T$). Athey et al. (2021, sec 3.1) discuss how the dimensions of the matrix lead to different advisable estimation strategies: when the outcome matrix is 'thin', outcome modelling strategies such as fixed effects or matching classified as 'horizontal regression'[30] can reconstruct the missing potential outcomes well. In contrast, when $N << T$, a 'vertical regression'[31] approach such as synthetic controls is better suited to imputing potential outcomes, since the small number of units necessitates learning the relationship between the treatment units and control units outcomes in the pre-period and extrapolating it forward. For approximately square outcome matrices when $N \approx T$, a mixture of the two approaches, called factor modeling or matrix factorization, is suitable.

Panel Match In our setting, $N \approx 30,000 >> T = 22$, and therefore a matching strategy is the most advisable, since we have sufficient candidate untreated cells to match treated cells' trajectories as closely as possible and use their average to impute the missing potential outcome.

Imai, Kim, and Wang (2021) propose a matching procedure that first matches on treatment trajectory. That is, for a unit treated at time t, find all units that have the same treatment history for periods t-k to t-1, which is challenging in the presence of general patterns of treatment assignment that includes reversals as in their democratization and growth case. This is not a problem in our case: since treatment assignment is absorbing (that is, once treated a pixel cannot revert to a control condition) and fixed at a point time at the state level (for example, all units in Jharkhand get treated in 2010 and stay treated after), all treated units initially get matched with all untreated units within that state because their treatment histories are the same. This is because of the within state comparison we impose with the exact match and fixed effects. So, it is impossible for units to have 0 matches as long as there exist untreated units within any given state and mining conflict histories. Furthermore, in the fixed-effects specifications, all untreated units get equal weights (Athey et al. 2021), which might be an undesirable assumption, and the particular way that Imai, Kim, and Wang (2021) 'refine' these matches is by modeling the treatment assignment using, for instance, propensity score/ covariate balancing propensity score. These are the two refinements we employ as well.

[30]because the rows of the **Y** matrix are the units of the regression

[31]because the columns of the matrix **Y** are the units of the regression

Interactive-Fixed Effects For completeness, we also estimate an 'interactive-fixed effects' model (Bai 2009). The regression is now written with the additional term $\lambda_i' f_t$ in equation 1, where λ_i are a $r-$ dimensional time varying confounder:

$$Y_{ist} = \tau \text{Scheduled Area}_{is} \times \text{PESA Election Year}_{ist} + \delta_i + \gamma_t + \lambda_i' f_t + \epsilon_{ist} \tag{3}$$

Here, the time-varying confound r is expected to be low-rank and can be computed using out-of-sample predictive accuracy (Xu 2017). When we estimate the above regression using our analysis sample from Table 1, we find that a $r = 1$ interactive fixed effect estimator maximises fit, and estimate an ATT of 0.185 with a standard error of 0.03. The results remain largely unchaged for values of r between 1 and 4.

Given the large scale of our data and the fact that we have multiple treatment units, synthetic control strategies such as Abadie, Diamond, and Hainmueller (2010) and Ben-Michael, Feller, and Rothstein (2021) were infeasible.

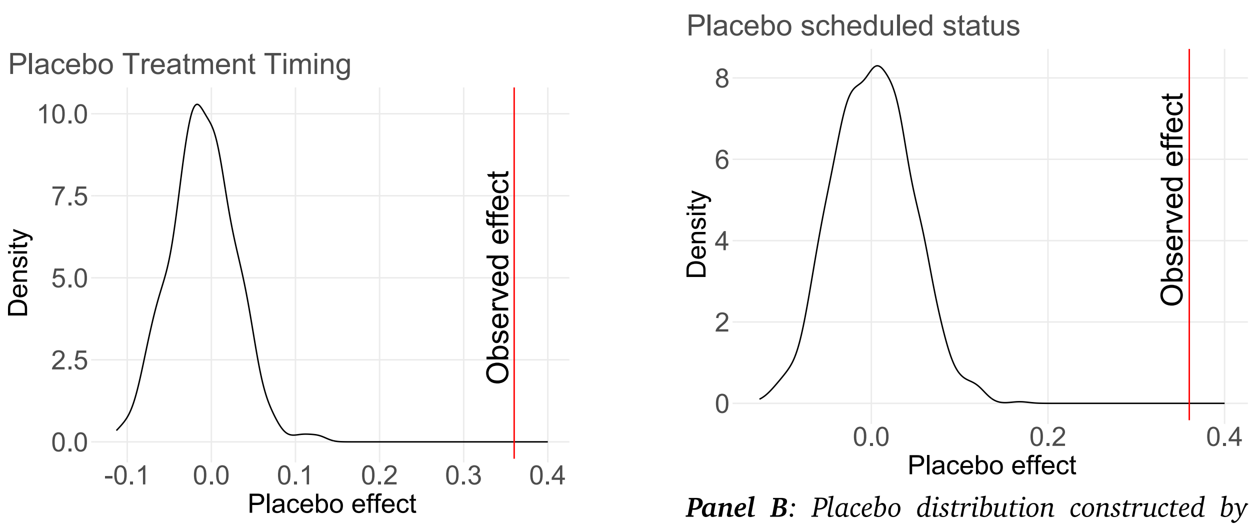
D.2 Placebo Regressions

We also conduct two placebo analyses motivated by the fact that treatment status D_{ist} in our design is defined as Scheduled status of cell i (illustrated in Figure 1) $\times$ time period $t \geq$ implementation of PESA within each state PESA$_s$ (illustrated in Figure 2). First, we hold a cell's Scheduled status fixed and randomly draw treatment timing from pre-treatment period years. That is, we sample a year from the light-blue section in Figure 2 for each state. We then re-compute our preferred specification to construct null the distribution in panel (A) of Figure A1. We restrict placebo treatment timing to pre-treatment years because, in the presence of dynamic treatment effects, placebo treatments assigned after the true treatment date (that is, placebos in the dark-blue region in Figure 2) will incorrectly absorb real treatment effects and falsely reject the null. We find that, as expected, our observed treatment effect is towards the right edge of the null distribution.

Second, we hold treatment timing within the state fixed, and permute Scheduled status *within* each state with probability equal to the actual share of Scheduled units in the state (that is, share of pixels inside gray cells within each state outlined in yellow in Figure 1). As before, we recompute our preferred specification to construct the null distribution in panel (B) of Figure A1. Again, we find that our observed effect is very large and lies to the right of the null distribution.

Taken together, these placebo estimates increase confidence in our findings and suggest that it is unlikely that our estimated effect is a type-1 error.

Figure A1: **Placebos**



***Panel A**: Placebo distribution constructed by holding cells' scheduled status fixed and permuting the timing of the treatment*

***Panel B**: Placebo distribution constructed by holding state-specific treatment timing fixed and permuting scheduled status (treatment eligibility) with probability equal to the share of scheduled units within the state*

D.3 Additional Results on the Role of Mandated Representation

Figure A2: **Roll-out of Local Government Institutions in Non-Scheduled villages following the 73rd Amendment**

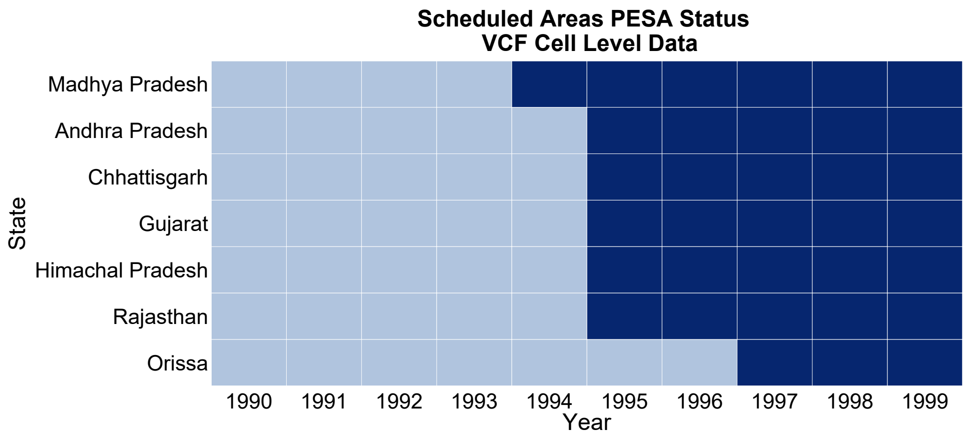
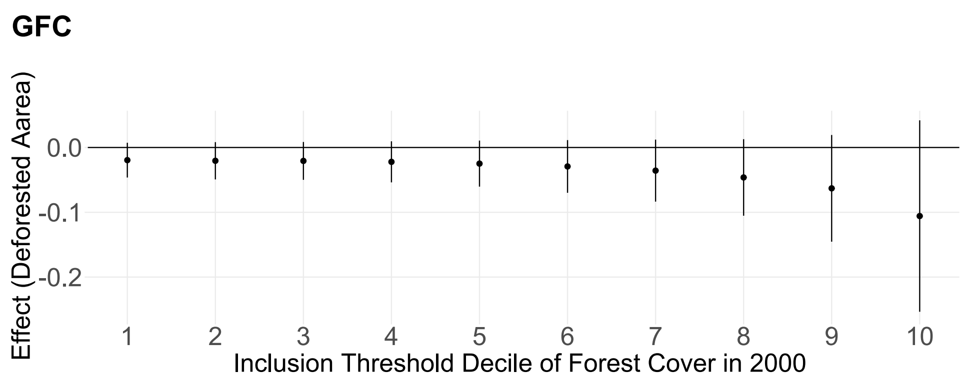


Figure A3: **Effects of Forest Rights act (2008) on deforested area in GFC data**



Notes: Standard errors are clustered by block. The replication materials also present these results in tabular form.

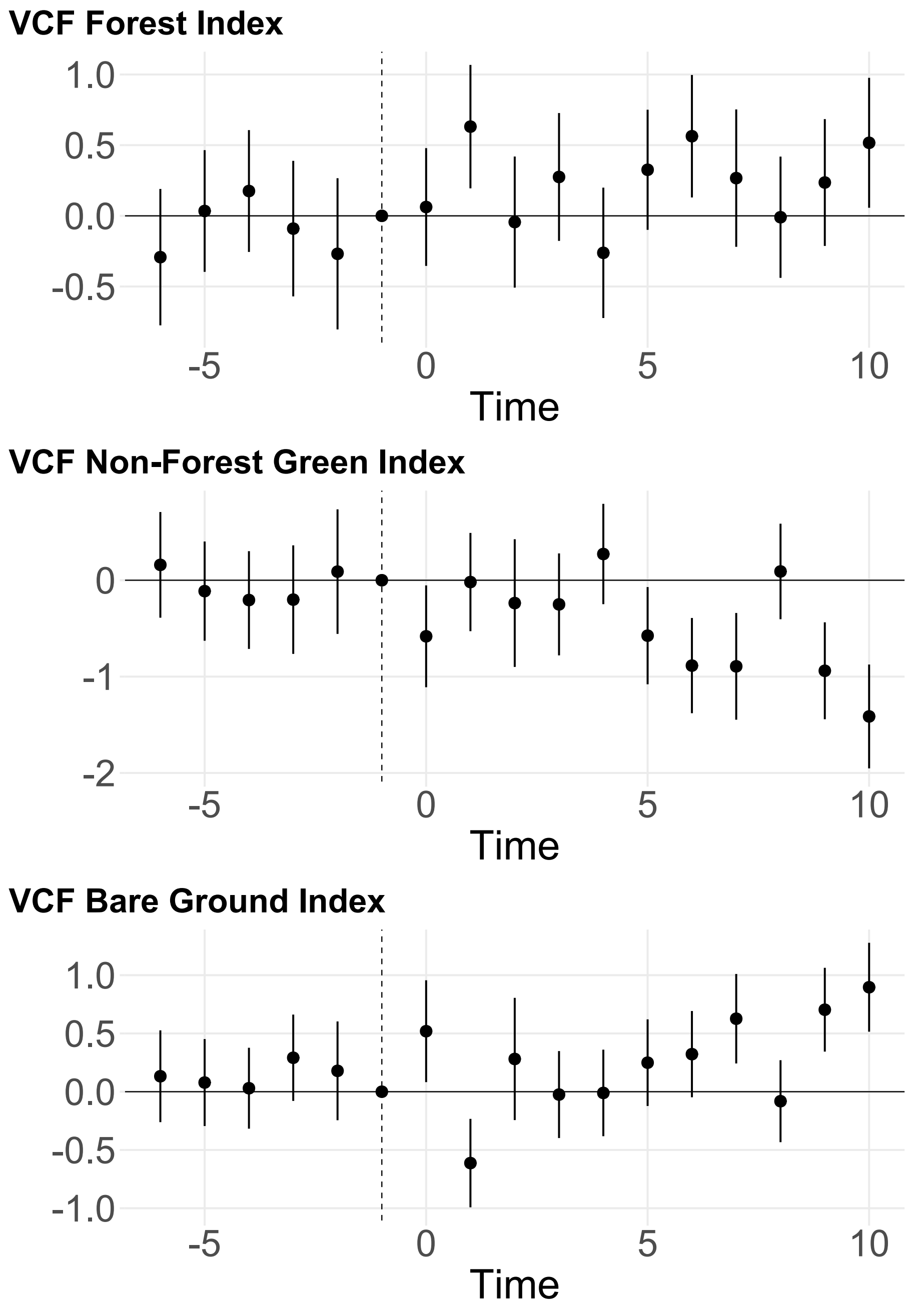
Table A5: Regression estimates decomposed by state (analysis sample at ex-ante median cutoff)

	Forest cover index			Annual Deforestation in Hectares		
	Full sample	Jharkhand	Others	Full sample	Jharkhand	Others
	(1)	(2)	(3)	(4)	(5)	(6)
PESA × Scheduled	0.3624	-0.0930	0.4122	-0.0672	-0.0816	-0.0631
	(0.1136)	(0.2088)	(0.1239)	(0.0328)	(0.0217)	(0.0418)
Summary Statistics						
Mean Pre-Y (Non-Sch)	8.809	8.588	8.826	0.0800	0.0500	0.0900
Mean Pre-Y (Sch)	12.30	9.292	13.14	0.1300	0.1000	0.1700
Dataset	VCF	VCF	VCF	GFC	GFC	GFC
Timespan	1995-2017	1995-2017	1995-2017	2001-2017	2001-2017	2001-2017
Fixed-effects						
Pixel	✓	✓	✓			
Village				✓	✓	✓
State × Year	✓	✓	✓	✓	✓	✓
Time Trends						
t (Pixel)	✓	✓	✓			
t (Village)				✓	✓	✓
Fit statistics						
# Pixel	30,843	1,876	28,967	–	–	–
# Village	–	–	–	31,601	5,015	26,586
# State × Year	198	22	176	68	17	51
# Observations	678,546	41,272	637,274	537,217	85,255	451,962
R^2	0.91565	0.88328	0.91684	0.42138	0.29522	0.42398

Notes: Standard errors are clustered at the block level and reported in parentheses.

D.4 Additional Results on Stewardship of the Forest

Figure A4: **Dynamic Treatment Effects of PESA Adoption on Forest Index, Non-Forest Vegetation, and Bare Ground Indices**



Notes: Standard errors are clustered by block. The replication materials also present these results in tabular form.

D.5 Additional Mining Results

This section reports additional results related to the *Opposing Commercial Interests* mechanism.

D.5.1 Mining Census Data

We report the distribution of mines from the mining atlas (accompanying Asher and Novosad (2021) in A5.

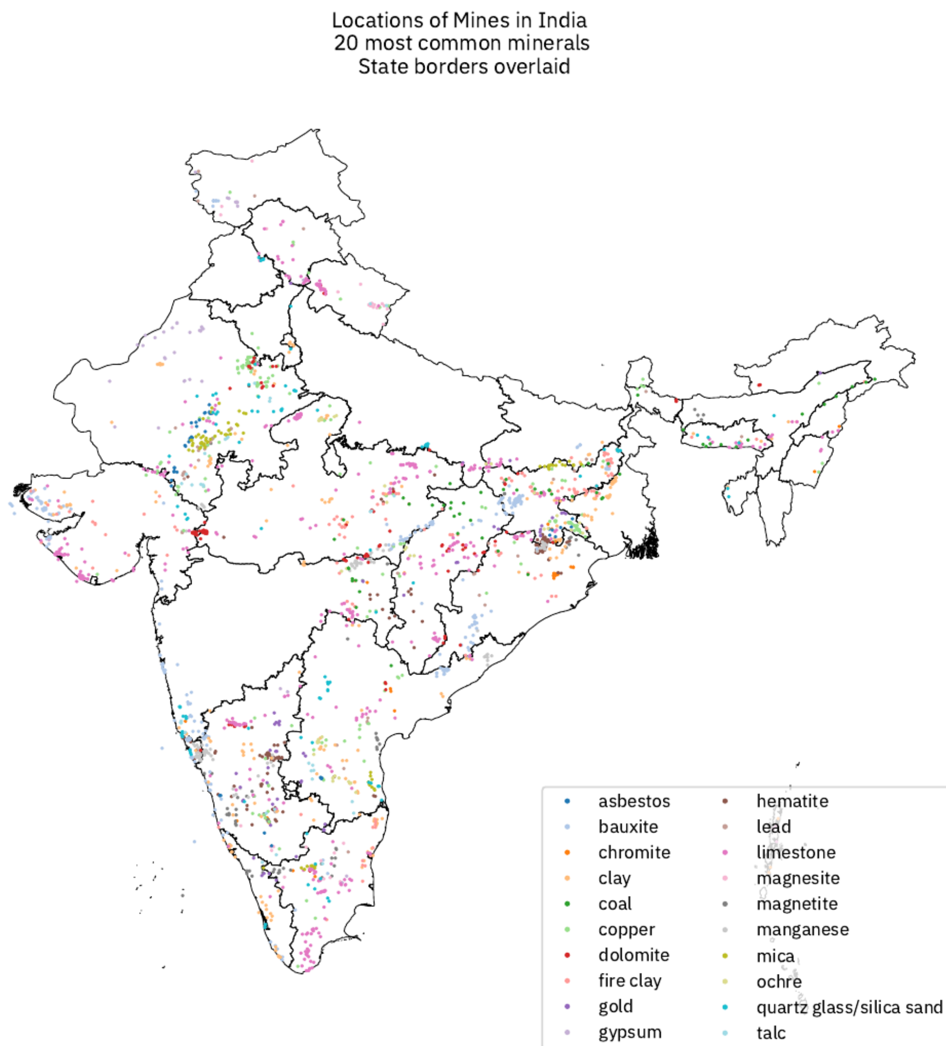


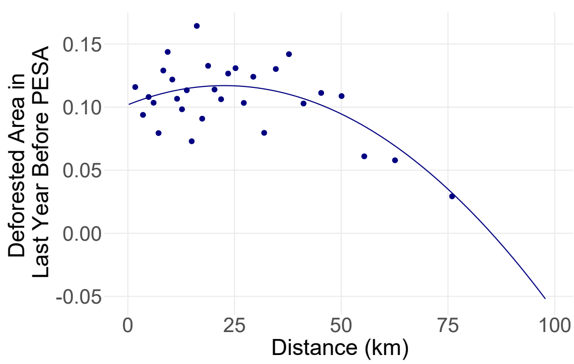
Figure A5: Location of 20 most common minerals

D.5.2 Robustness of VCF results in main text with GFC data

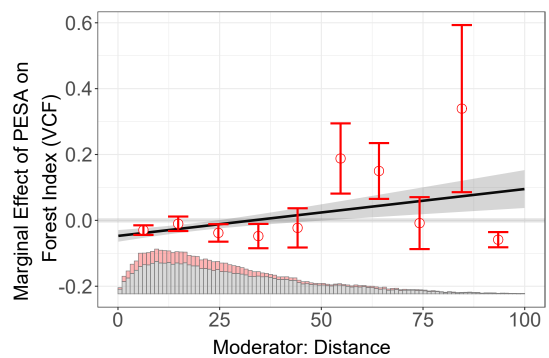
Here, we report deforested area in the year before PESA against distance to mines (panel A) and treatment effects moderated by distance to mines (panel B) using deforested-area in hectares from GFC, analogous to Figure 8 using VCF data. Panel A shows that, as with the VCF data, areas close to mines experience more deforestation prior to PESA. Similarly, as shown in Panel B, treatment effects on the reduction in deforestation are greatest in areas close to mines ex-ante.

Figure A6: **Deforestation and Proximity to Mining: GFC data**

***Panel A**: Pre-PESA Deforestation Rates (area in hectares) by distance to mines*



***Panel B**: Treatment effects on Annual Deforestation by Distance to Mines*



Notes for Panel A: The figure reports non-parametric binned scatterplot of decrease in forest index between 2001 and the first year before PESA as a function of distance to mines.
Notes for Panel B: We report treatment effects from a binned regression that estimates the treatment effect in pixels at different values of the moderator (distance to mines).
The replication materials also present these results in tabular form.

D.5.3 Mining proximity: Extensive and Intensive Margin

Next, we decompose the treatment effect by tercile of distance to the closest mine in Table A6, which uses VCF data in the first 3 column, and GFC data in the remaining 3 columns. The first tercile are areas closest to the mines, while the third tericle are those areas furthest from mines. The results show that areas that are closest to the mines ex-ante, have the largest effects on forest cover and deforestation.

In Table A7, column 1, we find that villages within 5 km of a mine experienced higher deforestation rates before PESA was implemented. In addition to studying treatment effect by proximity to *a* mine (the extensive margin), we also examine the effects of mining density (the intensive margin) in column 2. By intensive margin we mean the number of mines within a 5 km radius of each village in the GFC data.[32] We find that the treatment effect size grows within PESA villages as the number of mines close to the village increase. In column 2, we decompose this coarsely by estimating separate treatment effects for villages with 1-2, 3-4, and 5 or more mines within 5km, and find that the effects of PESA are increasing in the number of mines in close proximity to the village.

[32]We omit the analogous analysis for VCF because VCF pixels are much larger and therefore distance to mines are very noisily measured, thereby leaving us with little variation along the moderator (distance to mines).

Table A6: Regression estimates decomposed by distance to mines (ex-ante median cutoff)

	Forest cover index			Annual Deforestation in Hectares		
	(1)	(2)	(3)	(4)	(5)	(6)
Scheduled X PESA X 1st Tercile	0.1951	0.1373	0.2471	-0.0928	-0.0250	-0.0787
	(0.0705)	(0.0725)	(0.1344)	(0.0229)	(0.0213)	(0.0305)
Scheduled X PESA X 2nd Tercile	0.1318	0.1424	0.2814	-0.1025	-0.0261	-0.0771
	(0.0869)	(0.0773)	(0.1368)	(0.0277)	(0.0242)	(0.0394)
Scheduled X PESA X 3rd Tercile	-0.1256	-0.0017	0.3734	-0.0520	0.0185	-0.0309
	(0.1175)	(0.0996)	(0.1752)	(0.0413)	(0.0385)	(0.0675)
Summary Statistics						
Mean Pre-Y (Non-Sch)	7.046	7.046	7.046	0.0800	0.0800	0.0800
Mean Pre-Y (Sch)	10.76	10.76	10.76	0.1300	0.1300	0.1300
Dataset	VCF	VCF	VCF	GFC	GFC	GFC
Timespan	1995-2017	1995-2017	1995-2017	2001-2017	2001-2017	2001-2017
Fixed-effects						
Pixel	✓	✓	✓			
Village				✓	✓	✓
Year	✓			✓		
State × Year		✓	✓		✓	✓
Time Trends						
t (Pixel)			✓			
t (Village)						✓
Fit statistics						
# Pixel	41,449	41,449	41,449	–	–	–
# Village	–	–	–	52,776	52,776	52,776
# Year	22	–	–	17	–	–
# State × Year	–	198	198	–	153	153
# Observations	911,878	911,878	911,878	897,192	897,192	897,192
R^2	0.90889	0.91435	0.92081	0.35933	0.36780	0.41881

Notes: Standard errors are clustered at the block level and reported in parentheses.

Table A7: Regression estimates decomposed number of mines within 5km radius (ex-ante median cutoff)

	Annual Deforestation in Hectares	
	(1)	(2)
PESA × Scheduled	-0.0572	-0.0608
	(0.0339)	(0.0346)
PESA × Scheduled × # Mines in within 5km of village	-0.0464	
	(0.0263)	
PESA × Scheduled × Scheduled × 1-2 Mines in village		-0.0210
		(0.0334)
PESA × Scheduled × Scheduled × 3-4 Mines in village		-0.1594
		(0.0958)
PESA × Scheduled × Scheduled × 5+ Mines in village		-0.6187
		(0.3854)
Summary Statistics		
Mean Pre-Y (Non-Sch)	0.0800	0.0800
Mean Pre-Y (Sch)	0.1300	0.1300
Dataset	GFC	GFC
Timespan	2001-2017	2001-2017
Fixed-effects		
Village	✓	✓
State × Year	✓	✓
Time Trends		
t (Village)	✓	✓
Fit statistics		
# Village	52,776	52,776
# State × Year	153	153
# Observations	897,192	897,192
R^2	0.41882	0.41882

Notes: Standard errors are clustered at the block level and reported in parentheses.

Table A8: The Impact of Increased Representation on Mining Conflict Onset (Robustness)

	Mining Conflict Onset		
	30km	40km	50km
	(1)	(2)	(3)
PESA X Scheduled	0.0060	0.0101	0.0111
	(0.0037)	(0.0051)	(0.0063)
Mean Pre-Y (Non-Sch)	0.0133	0.0233	0.0361
Mean Pre-Y (Sched)	0.0344	0.0593	0.0854
Dataset	VCF	VCF	VCF
Timespan	1995-2017	1995-2017	1995-2017
Fixed-effects			
Pixel	✓	✓	✓
State × Year	✓	✓	✓
Time Trends			
t (Pixel)	✓	✓	✓
Fit statistics			
# Pixel	30843	30843	30843
# State × Year	198	198	198
# Observations	678546	678546	678546
Standard-Errors	Block	Block	Block
R^2	0.812	0.816	0.821

Notes: Standard errors are clustered at the block level and reported in parentheses. Conflict data is from `https://landconflictwatch.org/`. Conflict onset is assigned to all pixels within the indicated radius in each panel.

D.5.4 Robustness of Land Conflict Watch Results

Figure A7: **Effect of PESA on Mining Conflict Onset using `PanelMatch`**

Notes: The replication materials also present these results in tabular form.

E Software and Data used

The computation in this paper was performed in R and Python using the following libraries: Wickham (2010), Wickham et al. (2015), Gaure (2013), McKinney (2011), Jordahl (2014), Walt, Colbert, and Varoquaux (2011), Hunter (2007), Hainmueller, Mummolo, and Xu (2019), Cattaneo et al. (2019). Auxiliary data sources were Asher et al. (2020), Dimiceli et al. (2015), Asher and Novosad (2021), Hansen et al. (2013), Land Conflict Watch (2022)

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). “Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program”. *Journal of the American statistical Association* 105.490, pp. 493–505.

Asher, Sam and Paul Novosad (2021). “Rent-seeking and criminal politicians: Evidence from mining booms”. *Review of Economics and Statistics*.

Asher, Sam et al. (2020). “The Socioeconomic High-resolution Rural-Urban Geographic Dataset on India (SHRUG)”. World Bank Economic Review (Revise and Resubmit).

Athey, Susan et al. (2021). “Matrix completion methods for causal panel data models”. *Journal of the American Statistical Association* 116.536, pp. 1716–1730.

Bai, Jushan (2009). “Panel data models with interactive fixed effects”. *Econometrica* 77.4, pp. 1229–1279.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein (2021). “The augmented synthetic control method”. *Journal of the American Statistical Association* 116.536, pp. 1789–1803.

Béteille, André (1974). *Six Essays in Comparative Sociology*. New York: Oxford University Press.

Cattaneo, Matias D et al. (2019). “On binscatter”. *arXiv preprint arXiv:1902.09608*.

Corbridge, Stuart (2002). “The Continuing Struggle for India’s Jharkhand: Democracy, Decentralisation and the Politics of Names and Numbers”. *Commonwealth & Comparative Politics* 40.3, pp. 55–71.

Corbridge, Stuart, Sarah Jewitt, and Sanjay Kumar (2004). *Jharkhand: Environment, Development, Ethnicity*. Oxford University Press.

Dhebar, U N (1962). "Report of the Scheduled Castes and Scheduled Tribes Commission". *Government of India Press*.

Dimiceli, C et al. (2015). "MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250 m SIN Grid V006 [Data Set]". *NASA EOSDIS Land Process*.

Doudchenko, Nikolay and Guido W Imbens (2016). *Balancing, regression, difference-in-differences and synthetic control methods: A synthesis*. Tech. rep. National Bureau of Economic Research.

Ekka, Alexius and Mohammed Asif (2000). *Development-Induced Displacement and Rehabilitation in Jharkhand: A Database on its Extent and Nature*. Indian Social Institute.

Gadgil, Madhav and Ramachandra Guha (1992). "This fissured land".

Galanter, Marc (1984). *Competing Equalities : Law and the Backward Classes in India*. Oxford University Press.

Gaure, Simen (2013). "Lfe: Linear Group Fixed Effects". *The R Journal* 5.2, pp. 104–117.

Guha, Ramachandra (2000). *The Unquiet Woods: Ecological Change and Peasant Resistance in the Himalaya*. London: Oxford University Press.

Gulzar, Saad, Nicholas Haas, and Ben Pasquale (2020). "Representation and Development across Groups: Evidence from a Large-Scale Government Program in India". *American Political Science Review*.

Gupta, Sanjukta Das (2011). "Rethinking Adivasi Identity: The Chota Nagpur Tenancy Act (1908) and its aftermath among the Hos of Singbhum". *Adivasis In Colonial India: Survival, Resistance and Negotiation*. Ed. by Biswamoy Pati. Orient BlackSwan.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu (2019). "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice". *Political Analysis* 27.2, pp. 163–192.

Hansen, M C et al. (Nov. 2013). "High-resolution global maps of 21st-century forest cover change". *Science* 342.6160, pp. 850–853.

Hunter, John D. (2007). “Matplotlib: A 2D Graphics Environment”. *Computing in science & engineering* 9.3, pp. 90–95.

Imai, Kosuke, In Song Kim, and Erik H Wang (2021). “Matching Methods for Causal Inference with Time-Series Cross-Sectional Data”. *American Journal of Political Science*.

Jain, Meha (2020). “The benefits and pitfalls of using satellite data for causal inference”. *Review of Environmental Economics and Policy* 14.1, pp. 157–169.

Jordahl, K. (2014). “GeoPandas: Python Tools for Geographic Data”. *URL: https://github. com/geopandas/geopandas*.

Lancaster, Tony (2000). “The incidental parameter problem since 1948”. *Journal of econometrics* 95.2, pp. 391–413.

Land Conflict Watch (2022). “Conflicts Database”. *URL: https://www.landconflictwatch.org/about*.

McKinney, Wes (2011). “Pandas: A Foundational Python Library for Data Analysis and Statistics”. *Python for High Performance and Scientific Computing*, pp. 1–9.

Padel, Felix and Samarendra Das (2010). *Out of This Earth: East India Adivasis and the Aluminum Cartel*. Hyderabad: Orient: BlackSwan.

Patnaik, Sanjoy (2007). “PESA, the Forest Rights Act, and tribal rights in India”. *International Conference on Poverty Reduction and Forests. Bangkok*, pp. 3–7. URL: http://lib.icimod.org/record/13139/files/4946.pdf.

Shah, Alpa (2010). *In the Shadows of the State*. Duke University Press.

— (2013). “The Agrarian Question in a Maoist Guerrilla Zone: Land, Labour and Capital in the Forests and Hills of Jharkhand, India”. *Journal of Agrarian Change* 13.3, pp. 424–450.

Sharan, Ramesh (2009). “Forest Law, Ideology and Practice”. *Legal Grounds: Natural Resources, Identity and the Law in Jharkhand*. Ed. by Nandini Sundar. Oxford University Press.

Skaria, Ajay (1997). “Shades of wildness tribe, caste, and gender in Western India”. *Journal of Asian Studies*, pp. 726–745.

Song, Xiao-Peng et al. (Aug. 2018). “Global land change from 1982 to 2016”. *Nature* 560.7720, pp. 639–643.

Sundar, Nandini (2007). *Subalterns and Sovereigns: An Anthropological History of Bastar (1854-2006)*. London: Oxford University Press.

— (2009). "Laws Policies and Practices in Jharkhand". *Legal Grounds: Natural Resources, Identity, and the Law in Jharkhand*. Ed. by Nandini Sundar. Oxford University Press.

Vasan, Sudha (2009). "Forest Law, Ideology and Practice". *Legal Grounds: Natural Resources, Identity and the Law in Jharkhand*. Ed. by Nandini Sundar. Oxford University Press.

Verghese, Ajay (2016). "British Rule and Tribal Revolts in India: The curious case of Bastar". *Modern Asian Studies* 50.5, pp. 1619–1644.

Walt, Stéfan van der, S. Chris Colbert, and Gael Varoquaux (2011). "The NumPy Array: A Structure for Efficient Numerical Computation". *Computing in Science & Engineering* 13.2, pp. 22–30.

Wickham, Hadley (2010). "Ggplot2: Elegant Graphics for Data Analysis". *J Stat Softw* 35.1, pp. 65–88.

Wickham, Hadley et al. (2015). "Dplyr: A Grammar of Data Manipulation". *R package version 0.4* 3.

Xu, Yiqing (2017). "Generalized synthetic control method: Causal inference with interactive fixed effects models". *Political Analysis* 25.1, pp. 57–76.