Online Appendix
"Does Political Affirmative Action Work, and for Whom? Theory and Evidence on India's Scheduled Areas"

by Saad Gulzar, Nicholas Haas, and Benjamin Pasquale

A Background

A.1 Authors' Summary of Extant Work

Table A1: Authors' Summary of Past Work

Paper	Quota Targets	Quota Level	Effects on Targeted Minorities	Spillover Effects	Central Outcomes	# Observations
Dunning and Nilekani (2013)	SC/ST	Village	No	N/A	Reported participation in welfare programs	512 gram panchayats (GPs)
Jensenius (2015)	SC	State	No	None	Literacy, employment, village amenities	896 state constituences
Pande (2003)	SC/ST	State	Yes	Negative	Overall spending, spending on education, welfare	519 state-year
Chin and Prakash (2011)	SC/ST	State	Only ST	Positive	Poverty measures	627 state-year
Parthasarathy, Rao and Palaniswamy (2017)	Women	Village	Yes	Negative	Participation in conversation, state responsiveness	50 villages
Besley, Pande and Rao (2007)	SC/ST	Village	Yes	N/A	Government transfers to households	201 GPs (527 villages)
Duflo and Chattopadhyay (2004)	Women, SC/ST	Village	Yes	N/A	Constituent policy preferences, gov spending	265 GPs
Bardhan, Mookherjee and Torrado (2010)	SC/ST	Village	Yes	Positive	Household benefits: water, employment, etc	57 GPs (89 villages)
Palaniswamy and Krishnan (2012)	SC	Village	Yes	Positive	Spending	80 GPs (225 villages)
Bhavnani (2009)	Women	State	Yes	N/A	Electoral outcomes	118 election wards
Beaman et al. (2010)	Women	Village	Yes	None	Investment in drinking water	197 villages
Besley, Pande and Rao (2005)	SC	Village	Yes	N/A	Household beneficiary status	522 villages
Chauchard (2014)	SC	Village	Yes	N/A	Social attitudes, norms, stereotypes	64 GPs
Dunning (2010)	SC/ST	Village	Yes	N/A	Political attitudes and preferences	200 GPs

A.2   Role of Local Institutions in NREGS Implementation

The Panchayati Raj Institutions have a significant role to play in the implementation of a number of local development activities and program as shown in Figure A1. The degree and nature of their effort therefore is an important contribution to this significant variation across even local areas. See Figures A1 and A2 for details of the role of local government institutions in NREGS delivery.

Figure A1 shows that Panchayati Raj Institutions are responsible for planning, beneficiary selection, implementation, and monitoring and evaluation aspects of NREGS implementation. Similarly, Figure A1 shows that the Gram Panchayat is an important node for NREGS implementation that, nevertheless, rests in a wider networks of important agents.

Table 2: PRI Roles in Flagship Programmes

Sl.No.	Scheme	Planning	Beneficiary Selection	Implementation	Monitoring and Evaluation
1.	MGNREGS	✓	✓	✓	✓
2.	Backward Regional Grant Fund	✓		✓	✓
3.	Public Distribution System		✓		✓
4.	Shelter		✓	✓	
5.	Education		✓	✓	✓
6.	Health	✓		✓	✓
7.	Sanitation	✓	✓	✓	✓
8.	Electricity				✓
9.	Drinking Water	✓	✓	✓	✓
Source: India Rural Development Report 2012-13 by IDFC					

Figure A1: **PRI Duties (source: (ILO, 2015))**
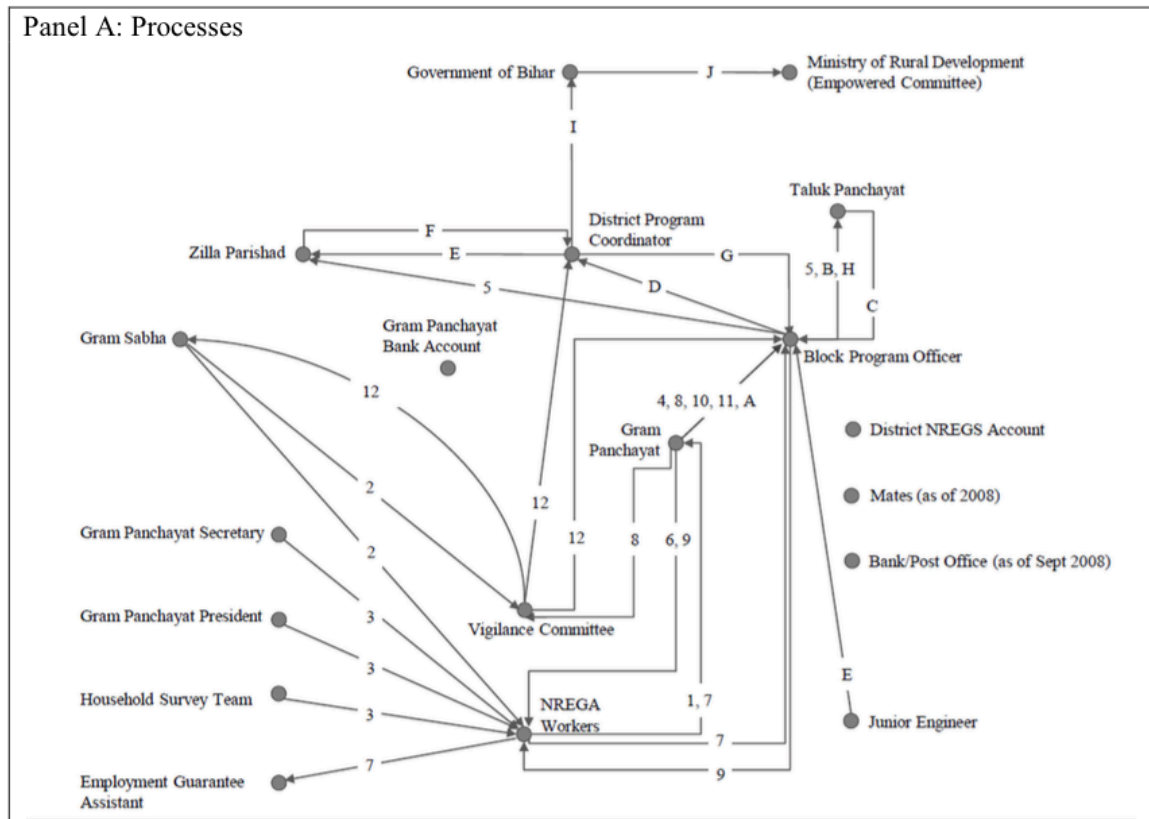
Panel A: Processes

Implementation process

1. Adult household members apply for registration.
2. Gram sabha verifies and mobilizes applications for registration and elects the members of the Vigilance and Monitoring Committee.
3. A household survey is performed for the identification of households that are willing to register for employment cards.
4. The gram panchayat sends a copy of the registration to the block program officer.
5. The block program officer reports to the taluk and zilla parishad.
6. The gram panchayat issues employment cards to registered households.
7. Adult household members apply for work at the gram panchayat or present themselves to the employment guarantee assistant (EGA). The EGA records the application in the employment register. The employment application via the block program officer is a "fallback" option only.
8. The gram panchayat informs (1) the block program officer about new work applications and (2) the Vigilance and Monitoring Committee about estimates regarding the work, time frame, and quality parameters.
9. The gram panchayat and block program officer assign employment.
10. The gram panchayat informs the block program officer about the employment allotments made.
11. The gram panchayat informs the block program officer about the start of work, and the block program officer issues registered Muster Rolls.
12. Upon completion of the work, the Vigilance and Monitoring Committee prepares the report and submits it to the gram sabha, the block program officer, and the district program coordinator.

Planning process

A. The gram panchayat forwards the development plan and priorities to the block program officer.
B. The block program officer scrutinizes and approves the gram panchayat–specific proposals and sends the consolidated gram panchayat proposals to the taluk panchayat.
C. The taluk panchayat sends the approval of the shelf of gram panchayat proposals.
D. The block program officer forwards the shelf of gram panchayat proposals to the district program coordinator.
E. The district program coordinator consolidates the block plans and prepares the labor budget. The junior engineer assists the block program officer and develops and approves technical estimates. These are sent to the zilla parishad, which approves the block-wise shelf of projects and the labor budget.
F. The zilla parishad informs the district program coordinator about the approval.
G. The district program coordinator reports the approved projects to the block program officer.
H. The block program officer forwards a copy of the block plan to the gram panchayats.
I. The district program coordinator forwards the labor budget to the Government of Bihar.
J. The Government of Bihar forwards the labor budget to the Ministry of Rural Development.

Figure A2: NREGS Implementation (source: Raabe, Sekher and Schiffer (2010))

A.3 Background on Political Quotas in India

Political quotas restrict the representation or leadership of government bodies. The government body may be a local or national administrative body, an elected local council, a state parliament or a national-level parliament.[21] While quotas most commonly target citizens (e.g. for positions as bureaucrats) or leadership positions (in councils or parliaments), quotas may also target particular geographic areas (e.g. with provisions for local autonomy).

Even prior to Indian Independence, the British government implemented quotas for individuals from particular ethnic identity categories. The Morley-Minto reforms in 1909 established separate electorates for Muslims. In 1919, the Mont-Ford reforms tied this 'quota' inversely to the proportion of Muslims in a given province (Rudolph and Rudolph, 2010, 560-561). Beyond political quotas for the national parliament, state parliaments, local government, Fifth Schedule Areas, several other types of political quotas exist in India. Broadly similar to the Fifth Schedule, the Indian Constitution's Sixth Schedule allows for the creation of Autonomous Councils. These councils for Scheduled Tribe communities, typically at the district or village levels, provide some legislative, administrative and judicial powers in areas now and formerly contained within the state of Assam (Chaudhury, 2005).[22]

Another example of a territorial quota is the delimitation of electoral boundaries. In both the national (*Lok Sabha*) and state (*Vidhan Sabha*) parliaments, the shape and number

---

[21]While not strictly meeting the definition of a type of government policy, political parties may also implement quotas. See for instance Mala Htun who analyzes why quotas for women are more likely to be utilized in parties but ethnic quotas more likely to be designed for legislatures (Htun, 2004).

[22]Reforms under the Sixth Schedule, also targeting the welfare of ST, according to our research, have been less consistent over time and space. Further we have been thus far unable to gather systematic data on the Autonomous District Councils at the heart of the Sixth Schedule, though we do believe this an area ripe for systematic research.

of electoral constituencies in a given area are determined by the size of the local population. Electoral redistricting was completed following every decennial census in 1952, 1963, 1973, and after a nearly three-decade delay, again in 2002 (Iyer and Reddy, 2013).[23] Even the linguistic reorganization of states, based on the States Reorganisation Act of 1956, can be considered a quota that generated state boundaries based on the relative linguistic homogeneity of a particular area (Tillin, 2013, Chapter 2).

For ordinary individuals, quotas influence individual's access to state education, government employment, and even rights to land. Reservations set aside places for individuals from the Scheduled Castes, Scheduled Tribes, women, in some states for individuals from Other Backward Classes, and even some religious groups (Corbridge, 2000; Galanter, 1984). Even land rights can be considered a type of quota. In Jharkhand for instance, customary laws such as the Chotanagpur Tenancy Act and the Santhal Parganas Tenancy Act both restrict land sales to individuals who are not associated with Scheduled Tribes (Sundar, 2005; Upadhya, 2009).

A.4   Background on Tribes in India

Early accounts of 'tribal' populations listed these groups as savages: animistic, violent, brutal, barbaric, wild. The *Fifth Report of the House of Commons* in 1812 described the inhabitants of Chotanagpur as a 'savage race, differing extremely in appearance, religion. British officials constructed Chotanagpur's 'tribes' through the lenses of 18th and 19th C. Victorian

---

[23]In 1977 delimitation was halted following complaints that delimitation according to population size incentivized certain population control policies. New legislation froze electoral boundaries in 1977 until delimitation was once again completed in 2002 Iyer and Reddy (2013, 5-6).

anthropology of racial types socio-cultural evolution (Damodaran, 2011, 58-59).[24] These administrator's conceptualizations, built not only on travelers' reports and racial theory but also their readings of sacred Hindu texts informed these constructions (Radhakrishna, 2011, 45-46).

British observers also made reference to the relative seclusion of Chotanagpur communities with references to the jungle, forest, inhospitable forest, wilderness and so-called 'primitive places.' S.C Roy wrote that Mundas settled in "primeval forests ... unmolested in their isolated mountain fastness ... walled off from the outside walls by chains of wooded hills" (1970: 60-61). As an example, anthropologist Verrier Elwin just after the end of the 19th C. proposed a system of national parks in order to preserve cultures of the Munda, Ho, Oraon and so on (Radhakrishna, 2011, 53).[25] Perceptions of Chotanagpur's 'tribes' roughly shifted from at first a wild savage to a 'noble' savage, and eventually to indigenous groups that needed to be protected from Aryan (then Hindu) invaders. In this way the British colonial government saw itself as "protectors of wild yet innocent tribals against rapacious outsiders" (Gupta, 2011a, 97). This shift followed the growth of 19th C. humanitarianism in Europe and the growth of missionaries in Chotanagpur.

---

[24]British officials such as Colonel Tickell and Ricketts in the 1840s-1850s, administrators W.W. Hunter and E. Dalton in late 19th C. and anthropologists S.C. Roy and Elwin Verrier around 1910, had all written of distinct Munda, Ho and Oraon communities. See (Damodaran, 2011; Gupta, 2011a,b; Galanter, 1984; Guha, 1996; Radhakrishna, 2011) for more on how British officials documented 'tribal' populations.

[25]Despite these accounts 19th C. reports of British officials and anthropologists make clear that communities in Singhbhum were not isolated but regularly interacted with groups in northern Jharkhand (Corbridge, Jewitt and Kumar, 2004).

A.5 Scheduled Tribes

India's 'tribal' identity category was first codified, with corresponding separate adminis-
trative areas specified, during the British Colonial period. Scholars have identified these
'tribal' groups (or *adivasi*) 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*a,b*). Despite regular mention of these factors, scholars agree that there has
been little clear definition or criteria as to what constitutes a 'tribe' (Béteille, 1974, 1986;
Dhebar, 1962; Corbridge, Jewitt and Kumar, 2004; Corbridge, 2002; Galanter, 1984).

Both prior to and following Indian Independence, leaders of the country have failed to
systematically define what constitutes a 'tribe' (or 'Scheduled Tribe'). Definitions that have
been given are vague, imprecise, and unclear – suggesting that lists of 'tribes' or ST were often
reflected the political convenience of whomever administered the region. Despite numerous
studies by Colonial administrator-anthropologist and close attention paid to the so-called
'tribes' of Chotanagpur, relatively little effort was given to writing rules for distinguishing
a tribe or tribal from the rest of the population.[26] British authorities first provided a list of
'Aboriginal Tribes' and 'Semi-Hinduised Aboriginal Tribes' in the Census of 1872 (Corbridge,
2002, 64). Census Commissioner H.H. Risley described a tribe as follows:

> A tribe as I find in India is a collection of families or groups of families bearing a
> common name which as a rule does not denote any specific occupation; generally
> claiming common descent from a mythical or historical ancestor and occasionally
> from an animal, but in some parts of the country held together by the obligations
> of blood-feud than by the tradition of kinship; usually speaking the same language

---

[26]Chotanagpur is a region of Eastern India covering parts of the states of Jharkhand,
Chhattisgarh, Odisha and West Bengal – a region with some of the largest tribal populations
in India and a region geographically proximate to the Colonial capital of Calcutta (today
Kolkata).

and occupying, professing, or claiming to occupy a definite tract of country. A tribe is not necessarily endogamous. (H.H. (1903, 514), as quoted in Pati (2011, 4).

In 1911, The Imperial Gazetteer of India provided a striking similar definition: "A collection of families bearing a common name, speaking a common dialect, occupying or professing to occupy a common territory and is not usually endogamous though originally it might have been so" (Nazer, 2004, 1). These definitions provide tremendous leeway for colonial officers to assign groups however they like. When J.H. Hutton, Indian Census Commissioner in 1931, sought to provide a list of 'tribes' he aimed to utilize the basis of "soul-substance" (Corbridge, Jewitt and Kumar, 2004, 30).

Despite the lack of clear criteria identifying so-called Tribals, special institutions were put in place for their protection with the Scheduled Districts Act of 1874. This territorial designation led to legislation with the aim of protecting tribals rights to their land, for instance through the Chotanagpur Tenancy Act of 1908. Upon Indian Independence from the British, the new constitution continued these policies of special administration in what were to become renamed "Scheduled Areas."[27]

In 1951 the First Report of the Commissioner for Scheduled Castes and Scheduled Tribes admitted that no precise method for identifying Scheduled Tribes had been created to date (Report 1951: 11). The report went on to note four characteristics identify a tribal: "tribal origin, primitive way of life, remote habitation, and general backwardness in all respects" (Report 1951: 109-111). Subsequent Commissions focused on Scheduled Castes and Scheduled Tribes reinforced the idea that little new information, methods of categorization or

---

[27]Under the Fifth Schedule of the Indian Constitution (1947), Scheduled Areas were created to allow customary practices and autonomy of Scheduled Tribes in these regions. By order of the President, a list of Scheduled Tribes and a list of Scheduled Areas was produced in 1950.

codification guided what constituted or defined either Scheduled Tribes or Scheduled Areas. Sociologist Andre Beteille has written, "lists of Indian tribes were in fact drawn up, with or without benefit of clear and consistent definitions" (Béteille, 1986, 299), and, "it cannot be too strongly emphasized that the list reflects the demands more of administrative and political circumstance than of academic or logical rigour" (Béteille, 1974, 62).

Indian Government officials even admitted the lack of definitions. According to the Dhebar Commission in 1961, "the term tribe is nowhere defined in the Constitution and in fact there is no satisfactory definition anywhere" (1962: 1). The Lokur Committee wrote when revising the list of Scheduled Tribes in 1965: "I have looked for indications of primitive traits, distinctive culture, geographic isolation, shyness of contact with the community at large and backwardness" (Galanter, 1984, 152).[28]

According to the Indian Constitution, Scheduled Areas are to define in those areas with a large fraction of the population belonging to a Scheduled Tribe. But this mapping of Scheduled Tribes to Scheduled Areas is equally unclear. Officially, according to the Fifth Schedule of the Constitution the President has the right to Schedule or De-schedule Areas and does so in consultation with Governors of Indian states. The Dhebar Commission of 1962 proposed a Scheduled Area be identified according to the following four, relatively vague, criteria.[29]

---

[28]Galanter notes this rough definition was reused in 1976: "this language is utilized almost verbatim by the Home Minister more than 10 years later in the debate about revision of the list" (1984: 152, fn. 143).

[29]In the mid-1970s the Twenty-Fourth Report of the Commissioner for Scheduled Castes and Scheduled Tribes proposed a clearer rule that areas with more than fifty-percent Scheduled Tribe population should be Scheduled Areas (Commissioner for Scheduled Castes and Scheduled Tribes, N.d., 117). But as we will show below, no such 50% threshold exists in terms of defining Scheduled Areas.

1. Preponderance of tribals in the population

2. Compact and reasonable size

3. Under-developed nature of the area

4. Marked disparity in economic standards of the people

While academic research and popular accounts discussing the Scheduled Areas have focused on political quotas, Scheduled Areas and PESA have several additional features intended to benefit ST. These include the establishment of state-level Tribes Advisory Councils intended to advocate on behalf of ST, restrictions over the sale of land, in an effort to prevent alienation of land by ST, as well as, within villages, reinforcing a given village's *gram sabha* (a sub-*gram panchayat* body), as an important unit of local decision-making.

Appendix Figure A3 below shows while the probability that an area is demarcated as a Scheduled Area is rising in the share of the ST population, there is no discontinuous jump in Scheduled Areas status where Scheduled Tribes constitute more than 50% of a village's total population. We see a similar continuity in the probability if we conduct an RD on whether the village has an ST plurality.

A.6 Case Study: The State of Jharkhand

A.6.1 Who is an ST?

After Independence the Government of Bihar made a clear break from historical laws giving uniform but distinct rights to the whole of Jharkhand. By means of the Scheduled Areas Order 1950 and the Scheduled Tribes Order 1951, the Government of Bihar effectively halved the amount of territory defined as Scheduled Areas and substantially shrunk the number of groups classified as ST. While the British Census of Chotanagpur in 1872 listed 31 aboriginal and 31 semi-aboriginal groups, the Scheduled Tribes Order of 1951 listed 30 such ST communities. According to Corbridge, according to the Census of 1951, "just 31.15% of
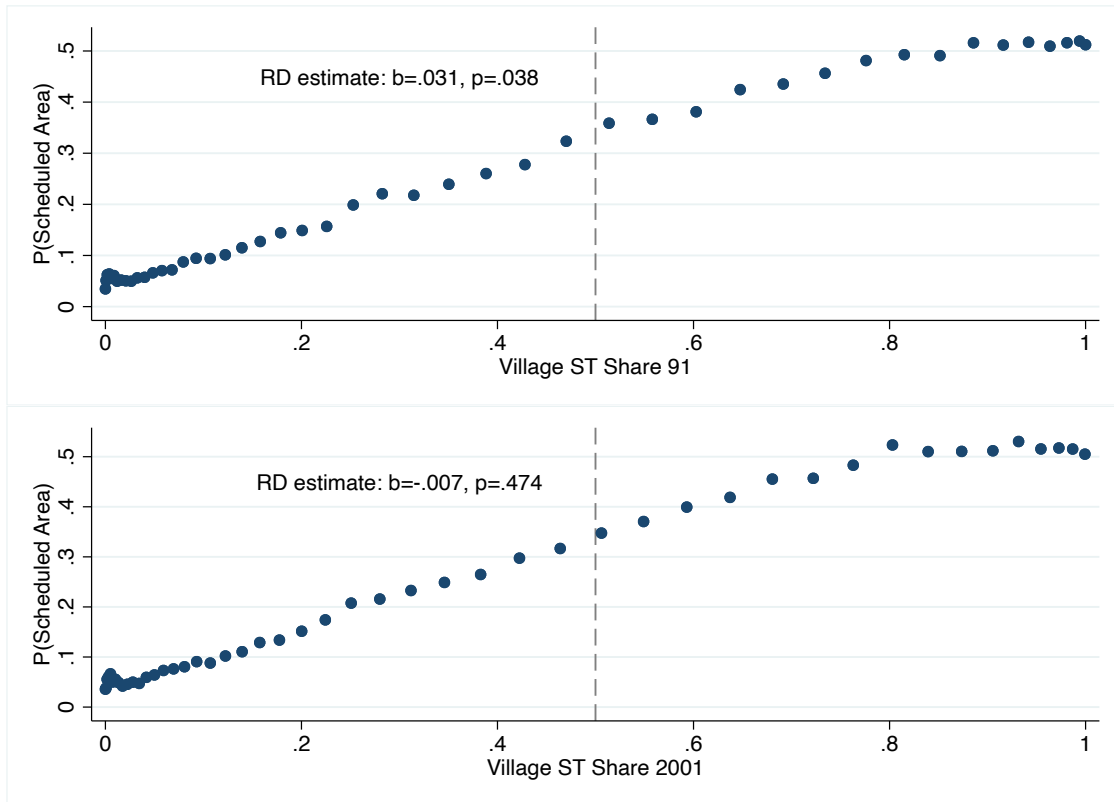
Figure A3: **Assignment of Scheduled Areas Status**

Notes: This figure shows how a village's ST population share (in 1991 and 2001) affects its likelihood of receiving Scheduled Areas status. Results attest to the haphazard assignment of Scheduled Areas, as they show that there is not a discontinuous jump in Scheduled Areas status where Scheduled Tribes constitute more than 50% of a village's total population.

the population of Chota Nagpur, and 44.67% of that of Santal Parganas was made up of Scheduled Tribals. Had the Census takers adopted the definitions used by the British in 1872, the percentage figures would have been 45.79% and 55.21% respectively" (Corbridge, Jewitt and Kumar, 2004, 64).[30]

A.6.2 Which Areas are Scheduled?

In practice, in Jharkhand today, most Scheduled Areas are assigned at the unit of district but some blocks are assigned as Scheduled Areas within Nonscheduled districts and some village-clusters are Scheduled within Nonscheduled blocks. With reference to earlier suggested criteria for which regions should be Scheduled: Jharkhand does not follow the fifty-percent rule as a criterion for Scheduling Areas. With no modifications in scheduling at the district-level, the Scheduled Areas assigned for Jharkhand (then Bihar) in 1950 have remained almost completely unchanged to present.[31] The Scheduled Areas of Jharkhand were re-affirmed after being assigned in 1950 in the Bihar Scheduled Areas Regulation of 1969 and again in 1977 and 2007.

---

[30]It is not exactly clear why certain groups were re-classified or if there was a clear methodology taken. Corbridge argues that some because some " 'aboriginals' had gained employment in the mining or industrial sectors was taken as evidence of their 'detribalisation"'' (Corbridge, Jewitt and Kumar, 2004, 64). Possibly the Government of Bihar thought de-scheduling some communities and areas would diminish the possibility that the mineral-rich region of Jharkhand would gain independent statehood.

[31]The only exceptions are the Scheduling of Bhandaria block of Garhwa district in 1977 and the Scheduling of two village-clusters, both within Satbarwa block in 2007.

B Data Construction

B.1 Creating an all-India dataset with NREGS, census, and election data sources

Because we hope this dataset and our procedures will be of use to other researchers we describe this process in detail:

1. Download the Socioeconomic High-resolution Rural-Urban Geographic Dataset for India (SHRUG), including available data from 1991-2011 Censuses and keys to match villages ($N \approx 647,000$) with 2001 and 2011 raw Census shape files (Asher and Novosad, 2019b).[32]

2. Download and combine village-cluster unit state datasets on NREGS from the MGN-REGA Public Data Portal.[33]

3. Build a village/village-cluster directory by downloading and combining individual block-level directory files from from the Ministry of Drinking Water and Sanitation.[34]

4. Extract and combine Census shape files using ArcGIS, to form spatially referenced (longitudes and latitudes) datasets of villages in the 2001 and 2011 Indian Censuses.[35]

---

[32]The SHRUG dataset may be accessed at: `https://doi.org/10.7910/DVN/DPESAK`.

[33]The MGNREGA Public Data Portal may be accessed at: `mnregaweb4.nic.in/netnrega/dynamic2/dynamicreport_new4.aspx`.

[34]We access the data from `https://web.archive.org/web/20110902154050/http://indiawater.gov.in/imisreports/nrdwpmain.aspx` at the National Rural Drinking Water Programme, Ministry of Drinking Water and Sanitation (Ministry of Drinking Water and Sanitation, 2014).

[35]We obtained Census data from New York University and Stanford University libraries, which licensed the data from InfoMap India (`https://www.mlinfomap.com/`).

5. Homogenize district and state names from NREGS dataset to the Water Ministry directory using a listing of all changes in district names and alternate spellings.[36] This allows matching of the NREGS dataset more efficiently.

6. Fuzzy match SHRID village names to the directory, and then NREGS village-cluster names to the directory. The directory provides a common reference for the two datasets. We used Stata's `reclink` command to carry out the fuzzy match, keeping the best available match. In some instances this was a perfect match as indicated by the matching score. In other instances, where the match was not perfect, we kept the best matches, while manually cleaning some of the data. Measurement error in matching is orthogonal to treatment as we show in Table A2.[37]

7. Add Scheduled Areas reservation status to the village dataset (see Appendix B.2 for more on Scheduled Areas reservation status construction).

8. Add assembly constituency-candidate level electoral records to the village dataset by locating each village within an assembly constituency using the village's latitude and longitude.[38]

9. Merge the dataset with Census 2001 and Census 2011 data using match keys from the SHRID dataset.

---

[36]For this, we rely on a compilation of all name changes between 2001 and 2011 available from (Statoids, N.d.), at `http://www.statoids.com/yin.html`.

[37]Other commands commonly used to fuzzy match string variables such as `soundex` are not useful in the Indian context because they rely on phonetic merging.

[38]Election data was downloaded from Election Commission of India (2014), at `https://web.archive.org/web/20141006084411/http://eci.nic.in/eci_main1/ElectionStatistics.aspx`. We used the Spatial Join command in ArcGIS to carry out this procedure.

Table A2: **Balance Table: Matched Villages Across India**

	Difference	p-value	Unmatched	Matched
Scheduled Areas	0.002	0.181	57457	216569
Population Index (Shrug, 1991-2001)	0.036	0.000	42765	213059
Minority Index (Shrug, 1991-2001)	-0.000	0.937	42765	213059
Public Goods Index (Shrug, 1991-2001)	0.116	0.000	42754	213048
Vulnerable Index (Shrug, 1991-2001)	-0.074	0.000	44197	213404

Notes: This table presents balance on variables that appear in both 1991 and 2001 Census waves between villages we are able to match in our dataset and those that remain unmatched. The 'Difference' column represents the effect of Matched on each outcome in rows. Importantly, matched villages are not more likely to be Scheduled Areas than are unmatched villages.

B.2 Identifying Scheduled Areas

Data on Scheduled Areas status was obtained from the government of India's Ministry of Tribal Affairs. The websites from which we obtained data in 2014 for eight of the nine states in our sample have since been retired, though they can be accessed today using Internet archive website The Wayback Machine. Below, we provide original links, as well as links that can still be used today to access the sites, for each of the eight states.

Andhra Pradesh

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinAndhraPradeshSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140818090711/http://tribal.nic.in:80/Content/ScheduledAreasinAndhraPradeshSSAreas.aspx`

Gujarat

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinGujarat.aspx`

- Archive link: `https://web.archive.org/web/20140818090722/http://tribal.nic.in:80/Content/ScheduledAreasinGujarat.aspx`

Jharkhand

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinBiharSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140818090717/http://tribal.nic.in:80/Content/ScheduledAreasinBiharSSAreas.aspx`

Himachal Pradesh

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinHimachalPradeshSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140818090727/http://tribal.nic.in:80/Content/ScheduledAreasinHimachalPradeshSSAreas.aspx`

Maharashtra

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinMaharashtraSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140818090843/http://tribal.nic.in:80/Content/ScheduledAreasinMaharashtraSSAreas.aspx`

Madhya Pradesh

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinMadhyaPradeshSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140818090732/http://tribal.nic.in:80/Content/ScheduledAreasinMadhyaPradeshSSAreas.aspx`

Odisha

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinOrissaSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140818090738/http://tribal.nic.in:80/Content/ScheduledAreasinOrissaSSAreas.aspx`

Rajasthan

- Original link: `http://tribal.nic.in/Content/ScheduledAreasinRajasthanSSAreas.aspx`

- Archive link: `https://web.archive.org/web/20140904021414/http://tribal.nic.in/Content/ScheduledAreasinRajasthanSSAreas.aspx`

Information on Scheduled Areas in all states, including the ninth in our sample, Chhattisgarh, may also be found in Annexure-II of "Statistical Profile of Scheduled Tribes in India (2013)," released by the Ministry of Tribal Affairs Statistical Division and accessible here: `https://tribal.nic.in/ST/StatisticalProfileofSTs2013.pdf`.

B.2.1 Verifying our Identification of Scheduled Areas and our spatial (longitudes and latitudes) data

.

To verify that we correctly identified Scheduled Areas, and more generally that our spatial (longitudes and latitudes) data are accurate, we can compare our map of Scheduled Areas that we generated using our data (Figure 2) to an official government map.[39] In Figure A4, we reproduce our map and compare it to the government map. We can see that our map closely matches the government map, but that ours provides more fine-grained information, bolstering our confidence in our data collection methods.

---

[39]This map can be accessed at `http://pesadarpan.gov.in/en_US/fifth-schedule-areas/-/asset_publisher`

Figure A4: Validating Spatial Data and Scheduled Area Identification

(a) Our Map
(Coding at Block Level)

(b) Government Map
(Coding at District Level)
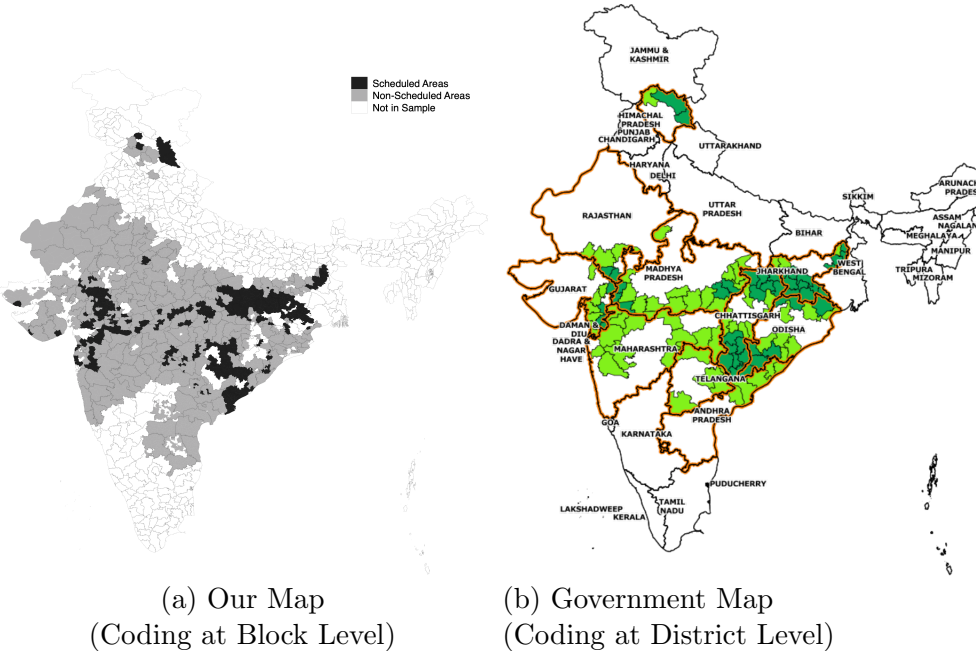
C   Summary Statistics, Balance, and Sorting

C.1   Optimal Bandwidths by Outcome

Table A3:   **Optimal Bandwidths by Outcome**

	Bandwidth	GPs
Jobcards	25.2	216042
Jobcards, ST	22.7	216042
Jobcards, SC	23.3	216042
Jobcards, non-SC/ST	35.2	216042
Worked HH	21.2	216042
Worked HH, ST	25.9	216042
Worked HH, SC	24.7	216042
Worked HH, non-SC/ST	19.2	216042
Logged (Workdays + 1)	22.0	216042
Logged (Workdays ST + 1)	27.3	216042
Logged (Workdays SC + 1)	23.7	216042
Logged (Workdays non-SC/ST + 1)	19.5	216042
Workdays, Women	21.1	216042
Workdays, Men	22.8	216042
Jobcards	22.3	216042
Jobcards, ST	21.0	216042
Jobcards, SC	27.7	216042
Jobcards, non-SC/ST	24.3	216042
Worked HH	20.6	216042
Worked HH, ST	24.2	216042
Worked HH, SC	21.8	216042
Worked HH, non-SC/ST	19.3	216042
Workdays	21.1	216042
Workdays - ST	24.1	216042
Workdays - SC	18.7	216042
Workdays - non-SC/ST	23.1	216042
Workdays, Women	23.9	216042
Workdays, Men	18.9	216042

Notes: This table presents optimal bandwidths by outcome variable. We include a geographic control function. Optimal bandwidths are calculated using the Stata package rdrobust (Calonico and Titiunik, 2017). The reported bandwidth is the "MSE-optimal point estimation using a common bandwidth on both sides of the cutoff" (Calonico and Titiunik (2017), p. 400). The estimation uses regularization methods, following Imbens and Kalyanaraman (2012).

C.2 Summary Statistics

Table A4: **Summary Statistics for 2001 and 1991 Censuses**

	(1) Non-Scheduled Areas		(2) Scheduled Areas		(3) Difference	
	mean	sd	mean	sd	b	t
Population Index (Census, 2001)	-0.06	0.89	-0.08	0.92	0.02*	(1.86)
Minority Index (Census, 2001)	-0.01	0.91	0.00	0.95	-0.02	(-1.58)
Vulnerable Pop Index (Census, 2001)	0.08	0.79	0.05	0.77	0.03***	(3.04)
Education Index (Census, 2001)	-0.05	0.85	-0.06	0.92	0.01	(0.94)
Medical Facilities Index (Census, 2001)	0.02	1.11	0.02	1.09	-0.01	(-0.54)
Water Index (Census, 2001)	0.04	0.89	0.31	0.92	-0.26***	(-25.75)
Communications Index (Census, 2001)	-0.07	0.98	-0.18	0.92	0.11***	(10.28)
Banking Index (Census, 2001)	0.00	1.20	-0.10	0.58	0.10***	(10.19)
Road Index (Census, 2001)	0.03	0.97	0.05	0.89	-0.03***	(-2.73)
Urbanization Index (Census, 2001)	0.06	1.48	0.22	0.57	-0.16***	(-13.57)
Irrigation Index (Census, 2001)	-0.12	0.68	-0.14	0.66	0.02***	(2.67)
Agricultural Worker Index (Census, 2001)	-0.02	0.96	-0.01	0.98	-0.01	(-0.88)
Marginal Worker Index (Census, 2001)	0.10	1.05	0.12	1.06	-0.02*	(-1.73)
Non-Agricultural Worker Index (Census, 2001)	-0.08	0.77	-0.10	0.77	0.01*	(1.71)
Share SC (Census, 1991)	0.09	0.14	0.07	0.12	0.02***	(16.35)
Share SC (Census, 2001)	0.09	0.14	0.07	0.12	0.02***	(15.46)
Share ST (Census, 1991)	0.48	0.39	0.58	0.37	-0.11***	(-25.31)
Share ST (Census, 2001)	0.47	0.38	0.58	0.37	-0.11***	(-25.20)
Village Longitude	3.18	0.19	3.18	0.19	-0.00	(-1.57)
Village Latitude	0.88	0.08	0.88	0.08	-0.00	(-0.56)
Bandwidth (in Kilometers)	4.87	2.87	4.36	2.79	0.51***	(16.33)
AC Reserved, ST	0.46	0.50	0.61	0.49	-0.15***	(-27.96)
ST Majority	0.49	0.50	0.61	0.49	-0.12***	(-22.09)
Observations	18375		14266		32641	

Notes: This table presents summary statistics on 2001 and 1991 Census variables for treated and untreated units using our 10 km geographic regression discontinuity specification on the same sample as in our main analyses. Column 1 presents the mean and standard deviation of variables in non-Scheduled Areas, column 2 presents the mean and standard deviation of variables in Scheduled Areas, and Column 3 presents the difference (column 1 - column 2) and t-statistic from a two-sample t test.

Table A5: **Summary Statistics for NREGS, Employment Outcomes**

	(1) non-Scheduled Areas		(2) Scheduled Areas		(3) Difference	
	mean	sd	mean	sd	b	t
Logged (Workdays + 1)	8.06	2.51	8.10	2.66	-0.04	(-1.52)
Logged (Workdays ST + 1)	6.34	3.16	6.86	3.15	-0.52***	(-14.77)
Logged (Workdays SC + 1)	4.84	3.12	4.66	3.15	0.18***	(5.14)
Logged (Workdays non-SC/ST + 1)	6.68	2.76	6.42	2.87	0.27***	(8.45)
Workdays	9748.16	10499.24	11467.71	15104.98	-1719.54***	(-11.60)
Workdays - ST	4306.59	6367.91	6851.69	12279.96	-2545.10***	(-22.52)
Workdays - SC	1259.99	2130.40	1139.16	2032.10	120.82***	(5.22)
Workdays - non-SC/ST	4181.59	6448.20	3476.85	5384.62	704.74***	(10.75)
Total Workers	5.85	1.05	5.83	1.06	0.02	(1.33)
Total Workers - Men	5.30	1.04	5.25	1.06	0.04***	(3.79)
Total Workers - Women	4.89	1.22	4.94	1.17	-0.05***	(-3.69)
Main Workers	5.15	1.51	5.07	1.59	0.07***	(4.25)
Main Workers - Men	4.78	1.45	4.70	1.53	0.09***	(5.24)
Main Workers - Women	3.76	1.74	3.75	1.77	0.02	(0.80)
Marginal Workers	4.20	1.71	4.24	1.69	-0.04**	(-2.08)
Marginal Workers - Men	3.28	1.62	3.30	1.60	-0.02	(-1.24)
Marginal Workers - Women	3.64	1.72	3.71	1.69	-0.07***	(-3.88)
Other Workers	3.01	1.58	2.87	1.64	0.15***	(8.07)
Other Workers - Men	2.72	1.60	2.56	1.65	0.16***	(8.69)
Other Workers - Women	1.80	1.25	1.76	1.26	0.04***	(2.77)
Observations	18375		14266		32641	

Notes: This table presents summary statistics on NREGS and 2011 Census employment outcome variables for treated and untreated units using our 10 km geographic regression discontinuity specification on the same sample as in our main analyses. Column 1 presents the mean and standard deviation of variables in non-Scheduled Areas, column 2 presents the mean and standard deviation of variables in Scheduled Areas, and Column 3 presents the difference (column 1 - column 2) and t-statistic from a two-sample t test.

Table A6: **Summary Statistics and Balance for 2011 Public Goods Analysis**

	(1) non-Scheduled Areas		(2) Scheduled Areas		(3) Difference	
	mean	sd	mean	sd	b	t
Population Index (Census, 2001)	0.77	1.69	1.01	1.96	-0.23***	(-3.25)
Minority Index (Census, 2001)	0.71	1.58	0.93	1.83	-0.23***	(-3.41)
Vulnerable Pop Index (Census, 2001)	0.19	0.76	0.11	0.79	0.08***	(2.76)
Education Index (Census, 2001)	0.64	1.69	0.81	1.96	-0.17**	(-2.36)
Medical Facilities Index (Census, 2001)	0.68	2.80	0.91	2.49	-0.23**	(-2.23)
Water Index (Census, 2001)	0.30	0.77	0.68	0.84	-0.39***	(-12.24)
Communications Index (Census, 2001)	0.44	1.40	0.50	1.47	-0.06	(-1.11)
Banking Index (Census, 2001)	0.41	1.52	0.46	1.29	-0.05	(-0.98)
Road Index (Census, 2001)	0.09	1.09	0.08	0.93	0.01	(0.24)
Urbanization Index (Census, 2001)	0.08	1.52	0.32	0.60	-0.23***	(-5.64)
Irrigation Index (Census, 2001)	0.21	0.96	0.29	1.12	-0.08**	(-2.07)
Agricultural Worker Index (Census, 2001)	0.76	1.60	1.01	1.74	-0.25***	(-3.86)
Marginal Worker Index (Census, 2001)	0.04	1.00	-0.01	0.91	0.06	(1.59)
Non-Agricultural Worker Index (Census, 2001)	0.45	1.54	0.55	1.54	-0.10*	(-1.66)
Share SC (Census, 1991)	0.11	0.12	0.09	0.10	0.02***	(5.11)
Share SC (Census, 2001)	0.11	0.12	0.09	0.10	0.02***	(4.47)
Share ST (Census, 1991)	0.35	0.34	0.44	0.36	-0.08***	(-6.01)
Share ST (Census, 2001)	0.36	0.34	0.44	0.35	-0.08***	(-6.20)
Observations	1632		1117		2749	

Notes: This table presents summary statistics on 2001 and 1991 Census variables for treated and untreated units using our 10 km geographic regression discontinuity specification on the same sample as in our analysis on 2011 public goods (Census market villages). Column 1 presents the mean and standard deviation of variables in non-Scheduled Areas, column 2 presents the mean and standard deviation of variables in Scheduled Areas, and Column 3 presents the difference (column 1 - column 2) and t-statistic from a two-sample t test.

Table A7: **Summary Statistics for 2011 Census Market Village Outcomes**

	(1) non-Scheduled Areas		(2) Scheduled Areas		(3) Difference	
	mean	sd	mean	sd	b	t
Overall Public Goods Index	0.51	0.21	0.55	0.19	-0.04***	(-5.52)
Roads Index	0.49	0.25	0.52	0.23	-0.04***	(-3.84)
All Weather Road (Y/N)	0.78	0.41	0.83	0.38	-0.04***	(-2.93)
National Highway (Y/N)	0.08	0.28	0.09	0.29	-0.01	(-0.75)
State Highway (Y/N)	0.18	0.38	0.21	0.41	-0.03**	(-2.04)
Major District Road (Y/N)	0.39	0.49	0.42	0.49	-0.03	(-1.62)
Other District Road (Y/N)	0.62	0.49	0.67	0.47	-0.05***	(-2.89)
Gravel Road (Y/N)	0.87	0.34	0.91	0.28	-0.04***	(-3.69)
Water Index	0.53	0.21	0.56	0.19	-0.03***	(-4.04)
Handpump (Y/N)	0.89	0.31	0.91	0.28	-0.02**	(-2.16)
Tap Water Treated (Y/N)	0.30	0.46	0.37	0.48	-0.07***	(-3.92)
Tank, Pond, Lake (Y/N)	0.49	0.50	0.48	0.50	0.02	(0.86)
Covered Well (Y/N)	0.16	0.37	0.21	0.41	-0.05***	(-3.26)
Uncovered Well (Y/N)	0.79	0.40	0.82	0.38	-0.03*	(-1.80)
Irrigation Index	0.39	0.36	0.41	0.36	-0.02	(-1.43)
Tubewell/Borehole (Y/N)	0.42	0.49	0.43	0.50	-0.02	(-0.79)
River/Canal (Y/N)	0.37	0.48	0.39	0.49	-0.03	(-1.33)
Electricity Index	0.70	0.38	0.72	0.36	-0.02*	(-1.71)
Electricity for Agriculture Use (Y/N)	0.67	0.47	0.71	0.46	-0.03*	(-1.87)
Electricity for Domestic Use (Y/N)	0.85	0.36	0.87	0.33	-0.02*	(-1.73)
Electricity for Commercial Use (Y/N)	0.57	0.49	0.59	0.49	-0.02	(-0.91)
Communications Index	0.46	0.34	0.54	0.34	-0.08***	(-6.20)
Post Office (Y/N)	0.33	0.47	0.43	0.49	-0.10***	(-5.37)
Post and Telegraph Office (Y/N)	0.24	0.43	0.32	0.47	-0.08***	(-4.45)
Telephone (Y/N)	0.53	0.50	0.60	0.49	-0.07***	(-3.67)
Mobile Coverage (Y/N)	0.75	0.43	0.83	0.38	-0.08***	(-5.00)
Education Index	0.47	0.26	0.53	0.25	-0.05***	(-5.36)
Primary School (Y/N)	0.90	0.30	0.95	0.22	-0.05***	(-4.71)
Middle School (Y/N)	0.74	0.44	0.82	0.38	-0.08***	(-4.95)
Secondary School (Y/N)	0.45	0.50	0.54	0.50	-0.09***	(-4.42)
Senior Secondary School (Y/N)	0.25	0.43	0.29	0.46	-0.05***	(-2.73)
College (Y/N)	0.03	0.18	0.04	0.19	-0.01	(-0.83)
Observations	1632		1117		2749	

Notes: This table presents summary statistics on 2011 public goods Census variables for treated and untreated units using our 10 km geographic regression discontinuity specification on the same sample as in our analysis on 2011 public goods (Census market villages). Column 1 presents the mean and standard deviation of variables in non-Scheduled Areas, column 2 presents the mean and standard deviation of variables in Scheduled Areas, and Column 3 presents the difference (column 1 - column 2) and t-statistic from a two-sample t test.

Table A8: **Summary Statistics and Balance for Full versus Market Village Samples**

	(1) Sample (Excl Market)		(2) Market Sample		(3) Difference	
	mean	sd	mean	sd	b	t
Population Index (Census, 2001)	-0.15	0.70	0.87	1.81	1.02***	(29.28)
Minority Index (Census, 2001)	-0.08	0.78	0.80	1.69	0.88***	(26.97)
Vulnerable Pop Index (Census, 2001)	0.06	0.78	0.16	0.78	0.10***	(6.19)
Education Index (Census, 2001)	-0.12	0.70	0.71	1.81	0.84***	(24.08)
Medical Facilities Index (Census, 2001)	-0.05	0.78	0.77	2.68	0.82***	(16.05)
Water Index (Census, 2001)	0.13	0.92	0.45	0.82	0.32***	(19.47)
Communications Index (Census, 2001)	-0.17	0.88	0.46	1.43	0.63***	(22.75)
Banking Index (Census, 2001)	-0.09	0.92	0.43	1.43	0.52***	(18.71)
Road Index (Census, 2001)	0.03	0.92	0.09	1.03	0.05***	(2.60)
Urbanization Index (Census, 2001)	0.13	1.17	0.18	1.23	0.05**	(1.98)
Irrigation Index (Census, 2001)	-0.17	0.61	0.24	1.03	0.41***	(20.53)
Agricultural Worker Index (Census, 2001)	-0.10	0.83	0.86	1.66	0.96***	(29.92)
Marginal Worker Index (Census, 2001)	0.11	1.06	0.02	0.96	-0.09***	(-4.76)
Non-Agricultural Worker Index (Census, 2001)	-0.14	0.63	0.49	1.54	0.63***	(21.28)
Share SC (Census, 1991)	0.08	0.13	0.10	0.11	0.02***	(9.55)
Share SC (Census, 2001)	0.08	0.13	0.10	0.11	0.02***	(9.91)
Share ST (Census, 1991)	0.54	0.38	0.39	0.35	-0.15***	(-20.92)
Share ST (Census, 2001)	0.53	0.38	0.39	0.35	-0.14***	(-20.26)
Village Longitude	3.18	0.19	3.22	0.19	0.04***	(10.37)
Village Latitude	0.88	0.08	0.87	0.09	-0.00*	(-1.74)
Bandwidth (in Kilometers)	4.63	2.85	4.84	2.77	0.21***	(3.83)
AC Reserved, ST	0.53	0.50	0.44	0.50	-0.10***	(-9.99)
ST Majority	0.55	0.50	0.39	0.49	-0.16***	(-16.74)
Observations	29892		2749		32641	

Table A9: **Summary Statistics for Assembly Constituency (AC) ST Reservation, Scheduled Areas**

	(1) non-ST Reserved		(2) ST Reserved		(3) Difference	
	mean	sd	mean	sd	b	t
Population Index (Census, 2001)	-0.09	1.04	-0.07	0.84	-0.02	(-1.30)
Minority Index (Census, 2001)	-0.12	0.97	0.08	0.93	-0.21***	(-12.58)
Vulnerable Pop Index (Census, 2001)	0.17	0.83	-0.02	0.72	0.18***	(13.55)
Education Index (Census, 2001)	-0.11	1.04	-0.03	0.84	-0.08***	(-5.06)
Medical Facilities Index (Census, 2001)	-0.01	1.28	0.05	0.95	-0.05***	(-2.75)
Water Index (Census, 2001)	0.37	0.81	0.26	0.98	0.11***	(7.26)
Communications Index (Census, 2001)	-0.22	0.97	-0.15	0.89	-0.07***	(-4.51)
Banking Index (Census, 2001)	-0.08	0.62	-0.11	0.56	0.03***	(2.75)
Road Index (Census, 2001)	0.04	0.99	0.06	0.81	-0.02	(-1.51)
Urbanization Index (Census, 2001)	0.06	0.49	0.33	0.60	-0.27***	(-29.51)
Irrigation Index (Census, 2001)	-0.17	0.69	-0.13	0.64	-0.05***	(-4.19)
Agricultural Worker Index (Census, 2001)	-0.12	0.94	0.06	1.00	-0.18***	(-10.84)
Marginal Worker Index (Census, 2001)	0.18	1.12	0.08	1.01	0.11***	(5.76)
Non-Agricultural Worker Index (Census, 2001)	-0.04	0.89	-0.13	0.68	0.09***	(6.70)
Share SC (Census, 1991)	0.09	0.15	0.05	0.10	0.04***	(17.35)
Share SC (Census, 2001)	0.09	0.15	0.05	0.09	0.04***	(19.85)
Share ST (Census, 1991)	0.42	0.39	0.68	0.33	-0.26***	(-41.56)
Share ST (Census, 2001)	0.42	0.38	0.68	0.32	-0.26***	(-42.32)
Village Longitude	3.27	0.21	3.13	0.16	0.14***	(42.89)
Village Latitude	0.88	0.09	0.87	0.07	0.01***	(6.78)
Bandwidth (in Kilometers)	4.39	2.78	4.34	2.80	0.06	(1.20)
Observations	5525		8741		14266	

Notes: This table presents summary statistics on 2001 and 1991 Census variables for Scheduled Areas villages with (ST Reserved) and without (non-ST Reserved) overlapping AC-level ST reservations using our 10 km geographic regression discontinuity specification on the same sample as in our analyses. Column 1 presents the mean and standard deviation of variables in non-ST reserved areas, column 2 presents the mean and standard deviation of variables in ST reserved areas, and Column 3 presents the difference (column 1 - column 2) and t-statistic from a two-sample t test.

Table A10: **Summary Statistics for Assembly Constituency (AC) ST Reservation, Non-Scheduled Areas**

| | (1) | | (2) | | (3) | |
| | non-ST Reserved | | ST Reserved | | Difference | |
	mean	sd	mean	sd	b	t
Population Index (Census, 2001)	0.00	1.00	-0.13	0.72	0.13***	(9.99)
Minority Index (Census, 2001)	-0.09	0.88	0.08	0.94	-0.18***	(-13.02)
Vulnerable Pop Index (Census, 2001)	0.08	0.84	0.08	0.72	0.01	(0.46)
Education Index (Census, 2001)	-0.05	0.87	-0.05	0.83	-0.00	(-0.22)
Medical Facilities Index (Census, 2001)	0.02	1.30	0.01	0.84	0.01	(0.38)
Water Index (Census, 2001)	0.06	0.86	0.03	0.93	0.03**	(2.17)
Communications Index (Census, 2001)	0.02	1.01	-0.18	0.93	0.20***	(14.11)
Banking Index (Census, 2001)	-0.03	0.67	0.04	1.61	-0.07***	(-3.84)
Road Index (Census, 2001)	-0.01	0.95	0.07	0.98	-0.09***	(-5.98)
Urbanization Index (Census, 2001)	-0.05	0.76	0.20	2.01	-0.25***	(-10.77)
Irrigation Index (Census, 2001)	-0.08	0.66	-0.17	0.69	0.09***	(9.27)
Agricultural Worker Index (Census, 2001)	-0.04	0.97	0.00	0.95	-0.04**	(-2.55)
Marginal Worker Index (Census, 2001)	0.08	1.04	0.12	1.06	-0.04**	(-2.39)
Non-Agricultural Worker Index (Census, 2001)	-0.00	0.91	-0.18	0.56	0.17***	(15.79)
Share SC (Census, 1991)	0.11	0.15	0.07	0.13	0.05***	(22.33)
Share SC (Census, 2001)	0.11	0.15	0.06	0.12	0.05***	(22.49)
Share ST (Census, 1991)	0.30	0.34	0.68	0.33	-0.39***	(-77.89)
Share ST (Census, 2001)	0.30	0.33	0.68	0.33	-0.38***	(-77.62)
Village Longitude	3.21	0.20	3.14	0.16	0.06***	(23.48)
Village Latitude	0.88	0.09	0.87	0.07	0.01***	(8.81)
Bandwidth (in Kilometers)	4.96	2.87	4.77	2.86	0.20***	(4.69)
Observations	9937		8438		18375	

Notes: This table presents summary statistics on 2001 and 1991 Census variables for non-Scheduled Areas villages with (ST Reserved) and without (non-ST Reserved) overlapping AC-level ST reservations using our 10 km geographic regression discontinuity specification on the same sample as in our analyses. Column 1 presents the mean and standard deviation of variables in non-ST reserved areas, column 2 presents the mean and standard deviation of variables in ST reserved areas, and Column 3 presents the difference (column 1 - column 2) and t-statistic from a two-sample t test.

C.3 Balance

Table A11: **Balance Table - OLS with State FE**

	Difference	p-value	GPs	Villages
Population Index (Census, 2001)	-0.080	0.000	93875	206364
Minority Index (Census, 2001)	0.033	0.000	93875	206364
Vulnerable Pop Index (Census, 2001)	-0.066	0.000	93875	206364
Education Index (Census, 2001)	-0.025	0.000	93875	206364
Medical Facilities Index (Census, 2001)	0.008	0.189	93875	206364
Water Index (Census, 2001)	0.208	0.000	93875	206364
Communications Index (Census, 2001)	-0.196	0.000	93875	206364
Banking Index (Census, 2001)	-0.124	0.000	93875	206364
Road Index (Census, 2001)	-0.014	0.097	93875	206364
Urbanization Index (Census, 2001)	0.154	0.000	93875	206364
Irrigation Index (Census, 2001)	-0.182	0.000	93875	206364
Agricultural Worker Index (Census, 2001)	0.024	0.001	93875	206364
Marginal Worker Index (Census, 2001)	0.098	0.000	93875	206364
Non-Agricultural Worker Index (Census, 2001)	-0.075	0.000	93875	206364
Share SC (Census, 1991)	-0.098	0.000	93875	206364
Share SC (Census, 2001)	-0.100	0.000	93875	206364
Share ST (Census, 1991)	0.431	0.000	93875	206364
Share ST (Census, 2001)	0.427	0.000	93875	206364

Notes: This table presents balance between treated and untreated units using our OLS specification on the same sample as our OLS analysis (see Appendix D.1). Standard errors are clustered at the gram panchayat (GP) level. Controls include state fixed effects. The 'Difference' column presents the treatment effect of Scheduled Areas on each Index in rows.

Table A12: **Balance Table - 10 km RD with State FE**

	Difference	p-value	GPs	Villages
Population Index (Census, 2001)	-0.033	0.003	14933	32641
Minority Index (Census, 2001)	0.008	0.490	14933	32641
Vulnerable Pop Index (Census, 2001)	-0.024	0.028	14933	32641
Education Index (Census, 2001)	-0.002	0.873	14933	32641
Medical Facilities Index (Census, 2001)	0.008	0.555	14933	32641
Water Index (Census, 2001)	0.212	0.000	14933	32641
Communications Index (Census, 2001)	-0.100	0.000	14933	32641
Banking Index (Census, 2001)	-0.119	0.000	14933	32641
Road Index (Census, 2001)	0.000	0.993	14933	32641
Urbanization Index (Census, 2001)	0.153	0.000	14933	32641
Irrigation Index (Census, 2001)	-0.015	0.069	14933	32641
Agricultural Worker Index (Census, 2001)	-0.005	0.693	14933	32641
Marginal Worker Index (Census, 2001)	0.025	0.112	14933	32641
Non-Agricultural Worker Index (Census, 2001)	-0.014	0.126	14933	32641
Share SC (Census, 1991)	-0.023	0.000	14933	32641
Share SC (Census, 2001)	-0.021	0.000	14933	32641
Share ST (Census, 1991)	0.110	0.000	14933	32641
Share ST (Census, 2001)	0.109	0.000	14933	32641

Notes: This table presents balance between treated and untreated units using our 10 km geographic regression discontinuity specification on the same sample as in our main analyses. Standard errors are clustered at the gram panchayat (GP) level. Controls include state fixed effects and a geographic control function. The 'Difference' column presents the treatment effect of Scheduled Areas on each Index in rows.

Table A13: **Balance Table on 2001 Census for 2011 Market Villages - 10 km RD with State FE**

	Difference	p-value	GPs	Villages
Population Index (Census, 2001)	0.000	0.995	2293	2749
Minority Index (Census, 2001)	0.054	0.345	2293	2749
Vulnerable Pop Index (Census, 2001)	-0.042	0.150	2293	2749
Education Index (Census, 2001)	0.105	0.098	2293	2749
Medical Facilities Index (Census, 2001)	0.062	0.562	2293	2749
Water Index (Census, 2001)	0.307	0.000	2293	2749
Communications Index (Census, 2001)	-0.059	0.170	2293	2749
Banking Index (Census, 2001)	-0.072	0.157	2293	2749
Road Index (Census, 2001)	-0.040	0.333	2293	2749
Urbanization Index (Census, 2001)	0.197	0.000	2293	2749
Irrigation Index (Census, 2001)	-0.029	0.419	2293	2749
Agricultural Worker Index (Census, 2001)	0.000	0.993	2293	2749
Marginal Worker Index (Census, 2001)	0.030	0.385	2293	2749
Non-Agricultural Worker Index (Census, 2001)	0.038	0.470	2293	2749
Share SC (Census, 1991)	-0.018	0.000	2293	2749
Share SC (Census, 2001)	-0.015	0.000	2293	2749
Share ST (Census, 1991)	0.090	0.000	2293	2749
Share ST (Census, 2001)	0.090	0.000	2293	2749

Notes: This table presents balance between treated and untreated units using our 10 km geographic regression discontinuity specification on the same sample as in our analysis on 2011 public goods (Census market villages). Standard errors are clustered at the gram panchayat (GP) level. Controls include state fixed effects and a geographic control function. The 'Difference' column presents the treatment effect of Scheduled Areas on each Index in rows.
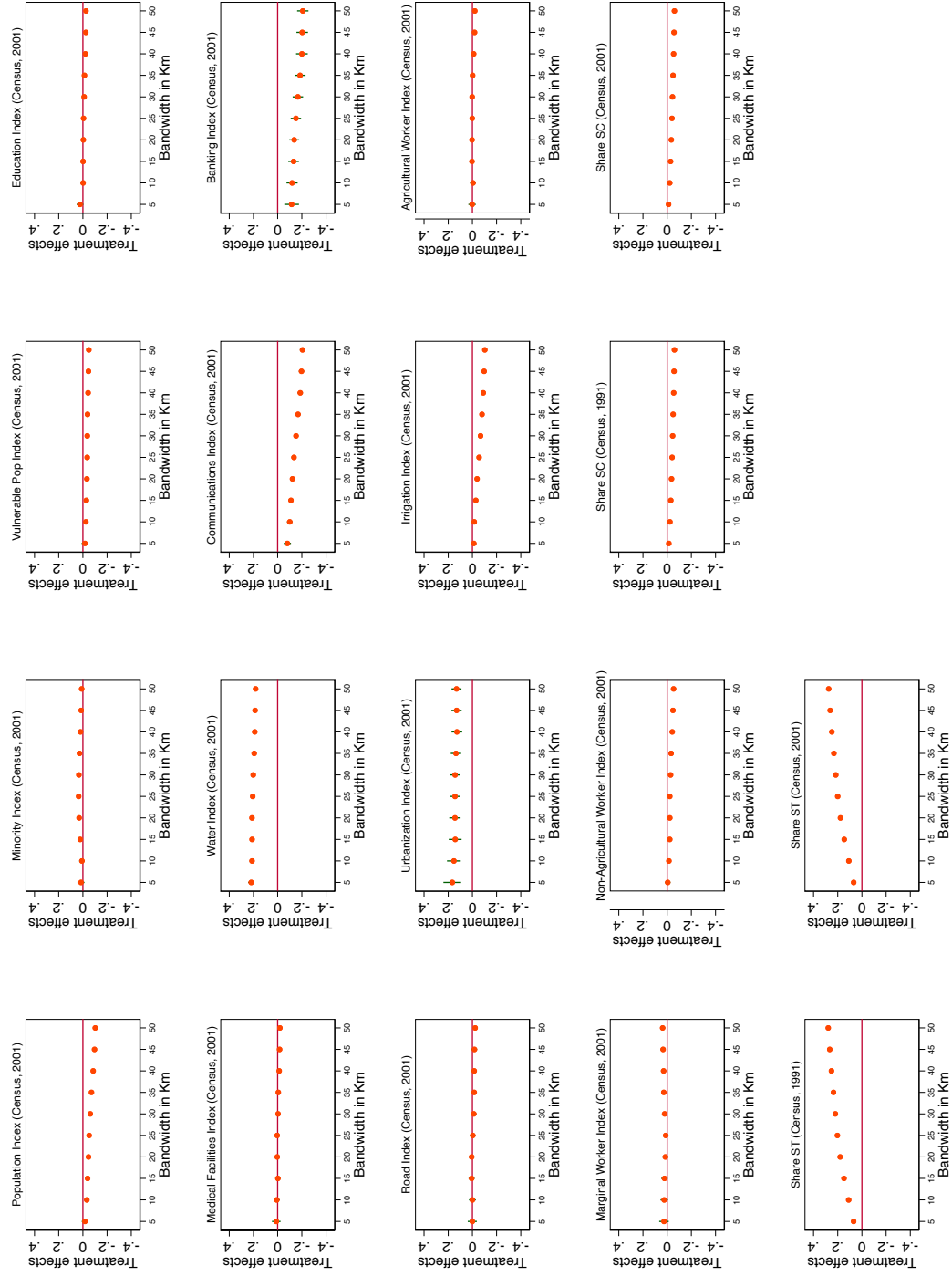
Figure A5: **Evolution of 1991 and 2001 Controls Across Different Bandwidths**

Notes: This figure shows the treatment effect of *Scheduled Areas* on all 14 indices (as well as 1991 and 2001 population shares) at various bandwidths. We find that the small imbalances that we do observe do not increase, but in most cases tend toward zero, as the bandwidth of analysis shrinks.

C.4 Over-time Changes and Sorting
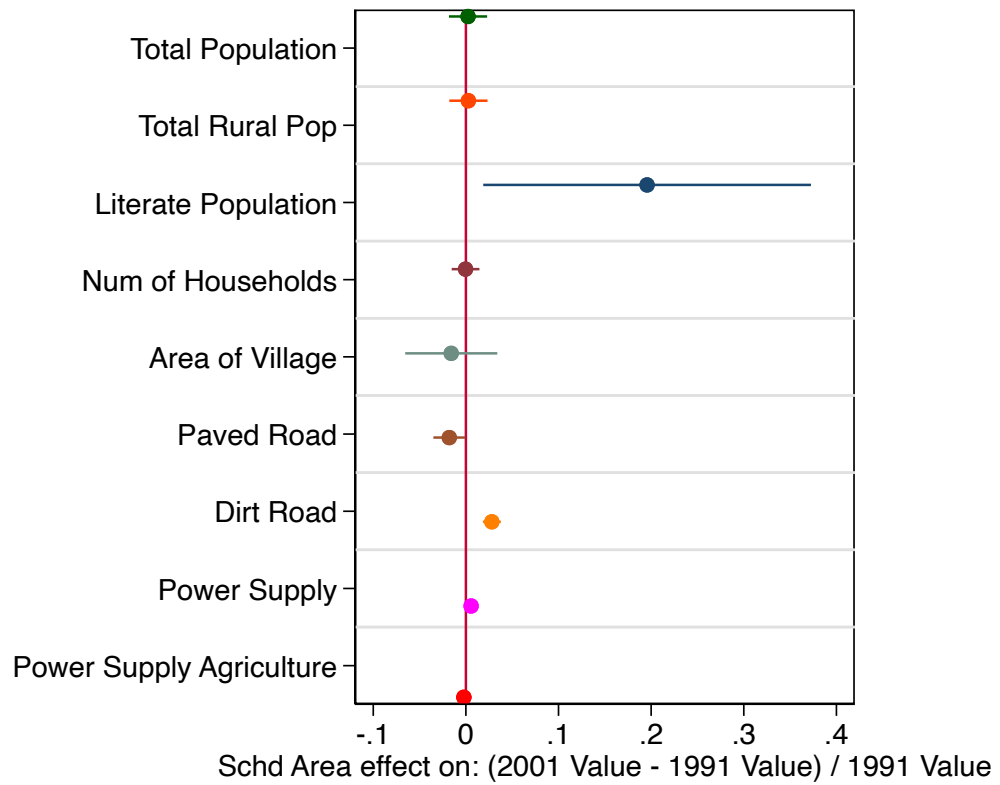
Figure A6: **Over-time Changes Across 1991-2001 Censuses**

Notes: This figure shows the treatment effect of *Scheduled Areas* on over-time changes in variables that we are able to track across both the 1991 and 2001 Censuses. We use our 10 km geographic regression discontinuity specification on the same sample as our main analysis. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion and state fixed effects. We find limited evidence of differential over-time changes across these variables in Scheduled versus non-Scheduled Areas.

Table A14: **Sorting Test**

	(1) Share ST 2001	(2) Share ST 2001	(3) Share ST 2011	(4) Share ST 2011	(5) Share ST 2011
Scheduled Areas	0.099***	0.005***	0.101***	0.007***	0.005***
	(0.006)	(0.001)	(0.006)	(0.001)	(0.001)
ST Share, 1991		0.941***			0.411***
		(0.002)			(0.023)
ST Share, 2001				0.953***	0.558***
				(0.002)	(0.023)
Control Mean	0.474	0.474	0.480	0.480	0.480
# GPs	14933	14933	14933	14933	14933
# Villages	32641	32641	32641	32641	32641

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. This table shows the treatment effect of *Scheduled Areas* on the ST share of the population in different Census years and controlling for different preceding years. We use our 10 km geographic regression discontinuity specification on the same sample as our main analysis. Standard errors are clustered at the gram panchayat (GP) level. Controls include state fixed effects, all 14 baseline indices, and a geographic control function. The table shows that once we account for ST population shares in 1991 and 2001, there remain no substantive differences in ST population across treated and control areas in 2011. We therefore control for 1991 and 2001 ST population shares in all of our reported analysis in the revised manuscript.
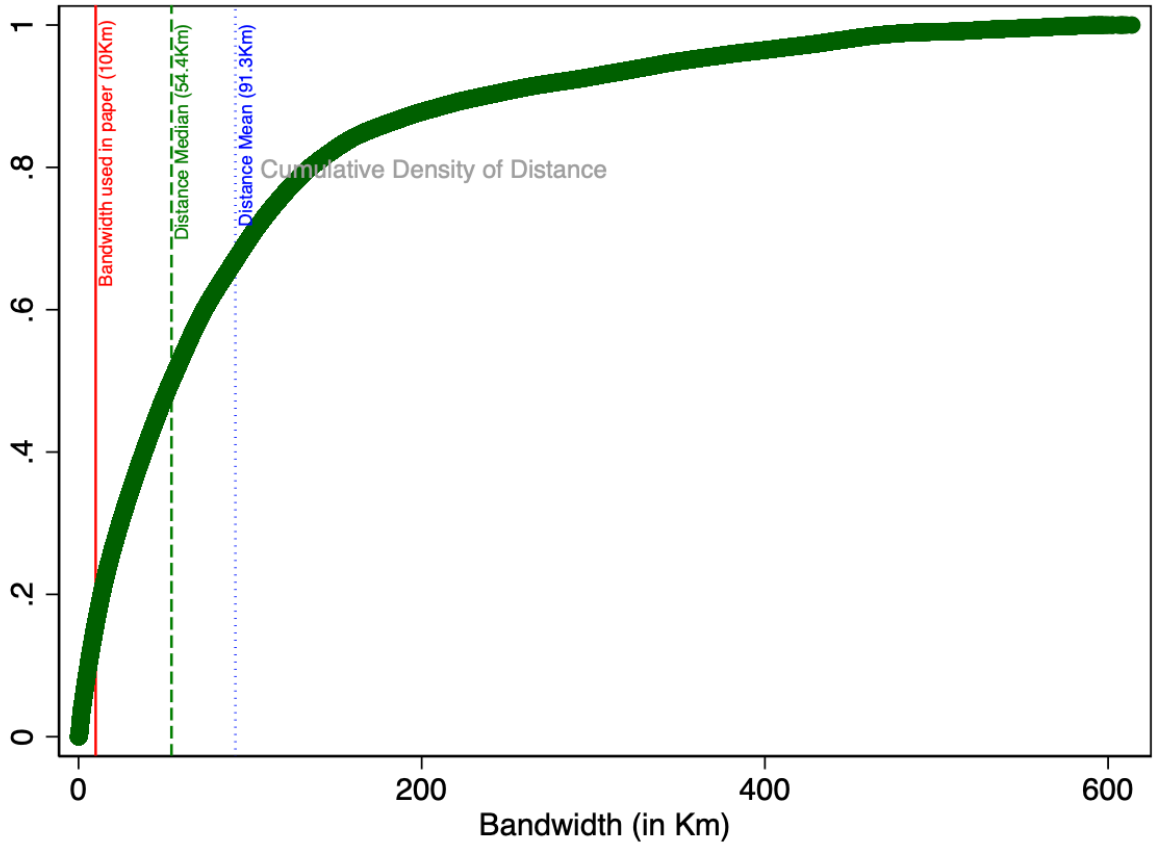
C.5 Density of Distance to Threshold

Figure A7: **Cumulative Density of Distance to the Threshold**

Notes: This figure plots the density of observations at different kilometer distances to the geographic regression discontinuity threshold. We include indicators for the bandwidth used in the paper (10 kilometers) and the mean and median distances to the threshold for reference.

D Robustness of Main Effects

D.1 OLS Main Results

A naive ordinary least squares model compares Scheduled and non-Scheduled villages with state fixed effects and village level controls from the 2001 and 1991 Indian Censuses with the following specification:

$$y_{vgs} \;=\; a_s + \gamma Scheduled\ Area_{vgs} + Z'_{vgs}\phi + \epsilon_{vgs} \tag{2}$$

Table A15: **The Effect of Scheduled Areas on NREGS (OLS)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	0.015*	0.594***	-0.439***	-0.339***
	(0.009)	(0.018)	(0.020)	(0.016)
Control Mean (Unlogged)	636.128	118.534	113.903	403.691
# GPs	93875	93875	93875	93875
# Villages	206364	206364	206364	206364
Panel B: Households Worked				
Scheduled Areas	0.109***	0.623***	-0.343***	-0.218***
	(0.014)	(0.019)	(0.021)	(0.019)
Control Mean (Unlogged)	192.633	41.400	37.695	113.538
# GPs	93875	93875	93875	93875
# Villages	206364	206364	206364	206364
Panel C: Workdays				
Scheduled Areas	0.127***	0.788***	-0.532***	-0.260***
	(0.022)	(0.031)	(0.035)	(0.029)
Control Mean (Unlogged)	8625.667	1795.946	1705.140	5124.582
# GPs	93875	93875	93875	93875
# Villages	206364	206364	206364	206364

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion and state fixed effects.

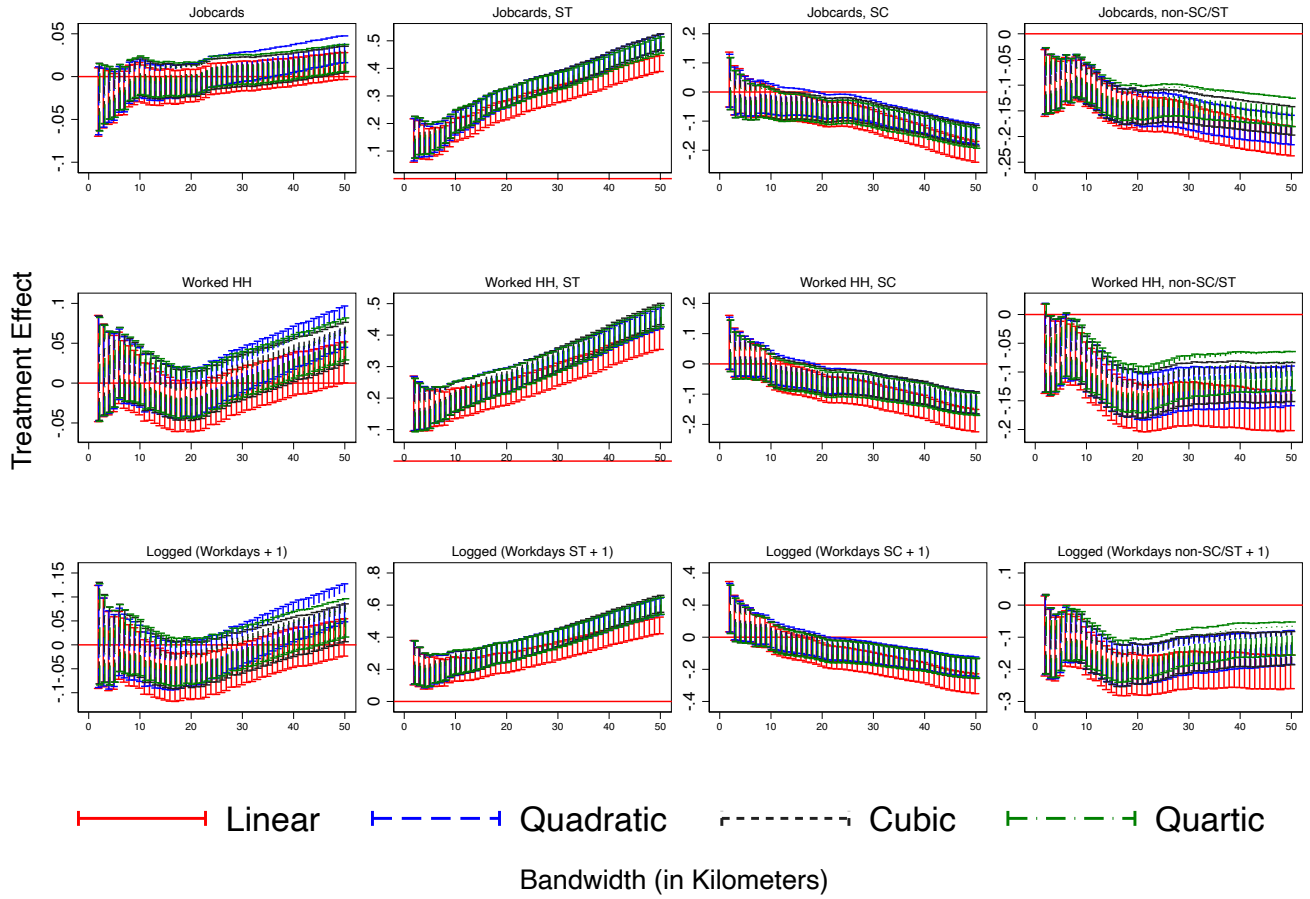## D.2   Functional Form and Bandwidth

Figure A8: **RD Robustness by Bandwidth and Functional Form**

Notes: Plots results in Table 2 by control function and bandwidth with 90% confidence intervals. Results show that coefficients reported as statistically significant in main table 2 are robust to various bandwidths and functional forms.

D.3 Controlling for Number of Matched Villages within a GP

One concern with the approach taken for the NREGS analysis could be that the number of villages within each gram panchayat differs between Scheduled and non-Scheduled Areas. Were this the case, then assigning all villages within a gram panchayat the same values could bias our comparison of Scheduled and non-Scheduled Areas. To account for this possibility, we show in table A16 that our main results are robust, both substantively and statistically, when we include fixed effects for the number of gram panchayat villages.

Table A16: **Main NREGS Results with Num Matched Villages per GP FE**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	-0.015	0.186***	-0.055*	-0.112***
	(0.013)	(0.024)	(0.030)	(0.022)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	-0.013	0.176***	-0.036	-0.109***
	(0.022)	(0.028)	(0.031)	(0.027)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	-0.034	0.206***	-0.017	-0.148***
	(0.035)	(0.045)	(0.053)	(0.043)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state and number of GP villages fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

D.4 Non-Logged Outcomes

Table A17: **The Effect of Scheduled Areas on NREGS (OLS)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	56.951***	140.602***	-28.336***	-55.314***
	(5.864)	(4.598)	(1.938)	(3.883)
Control Mean (Unlogged)	636.128	118.534	113.903	403.691
# GPs	93875	93875	93875	93875
# Villages	206364	206364	206364	206364
Panel B: Households Worked				
Scheduled Areas	58.593***	75.770***	-5.754***	-11.423***
	(3.332)	(2.625)	(0.812)	(1.702)
Control Mean (Unlogged)	192.633	41.400	37.695	113.538
# GPs	93875	93875	93875	93875
# Villages	206364	206364	206364	206364
Panel C: Workdays				
Scheduled Areas	4052.880***	4483.004***	-147.887***	-282.237***
	(230.799)	(175.919)	(41.439)	(97.659)
Control Mean (Unlogged)	8625.667	1795.946	1705.140	5124.582
# GPs	93875	93875	93875	93875
# Villages	206364	206364	206364	206364

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion and state fixed effects.

Table A18: **The Effect of Scheduled Areas on NREGS (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	19.990**	45.805***	-6.676**	-19.139***
	(7.999)	(6.249)	(2.808)	(5.679)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	23.389***	31.225***	-1.104	-6.732***
	(4.274)	(3.193)	(1.115)	(2.455)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	1671.018***	2117.170***	-49.564	-396.588***
	(270.954)	(203.103)	(49.948)	(128.846)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

Figure A9: **Leverage Plots, Logged and non-Logged Outcomes**

Notes: This figure displays leverage versus squared residual plots for logged and non-logged outcomes. The figure shows that residuals shrink under logged outcomes relative to non-logged outcomes. Consistent with these results, we observe improved goodness of fit and greater R² for our models that use logged outcomes.

D.5 Inverse Hyperbolic Sine Transformed Outcomes

Table A19: **The Effect of Scheduled Areas on NREGS (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	-0.002	0.213***	-0.049	-0.101***
	(0.014)	(0.027)	(0.034)	(0.025)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	0.001	0.212***	-0.024	-0.094***
	(0.025)	(0.032)	(0.035)	(0.032)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	-0.017	0.243***	-0.003	-0.127***
	(0.038)	(0.048)	(0.057)	(0.048)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. We transform outcomes with inverse hyperbolic sine transformation instead of the logarithmic transformation that we use throughout the paper. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2 y + xy^2$.
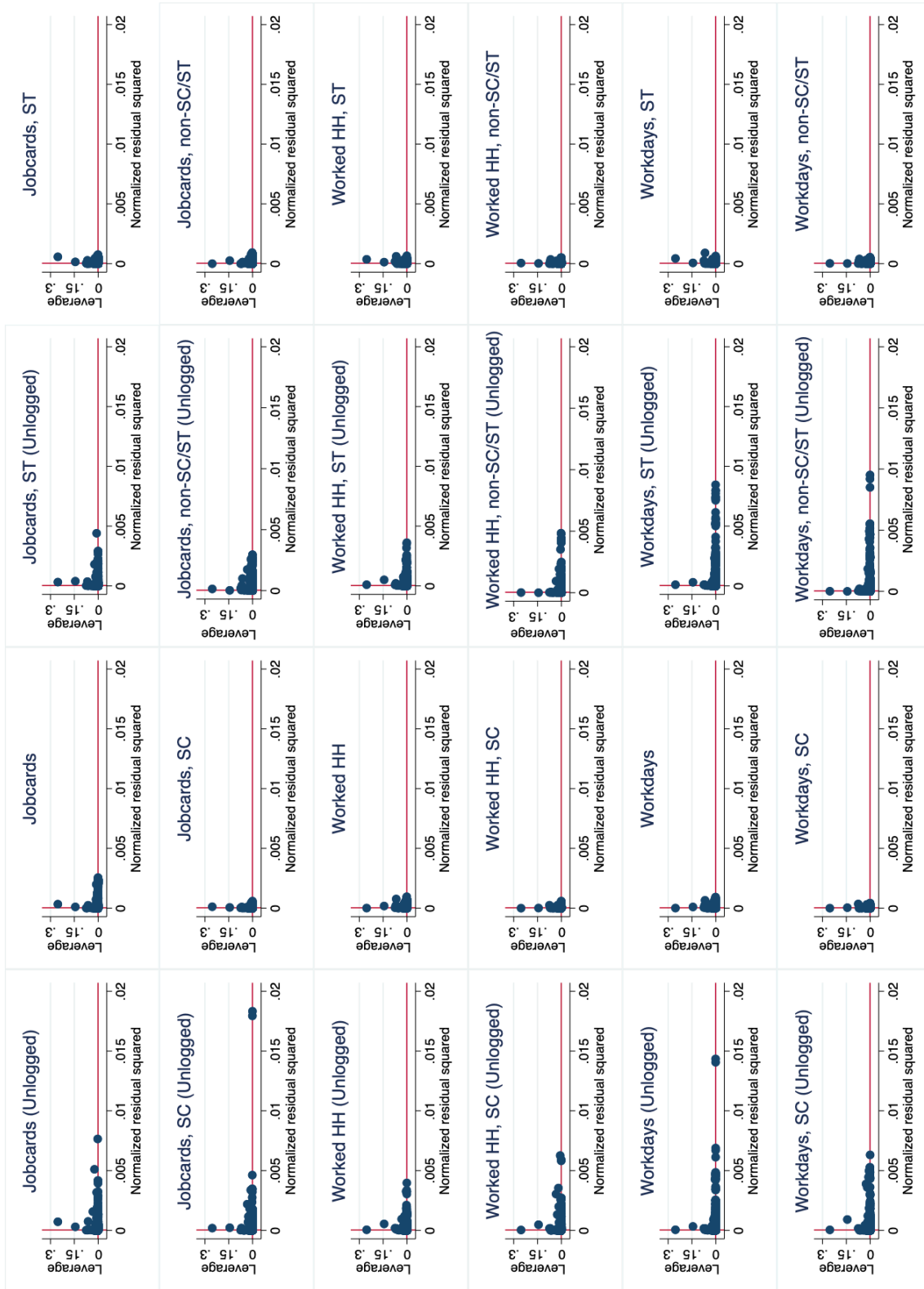
D.6 Standardized outcomes

Table A20: **The Effect of Scheduled Areas on NREGS (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	0.074***	0.297***	-0.081***	-0.152***
	(0.018)	(0.025)	(0.021)	(0.022)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	0.165***	0.354***	-0.033	-0.111***
	(0.022)	(0.028)	(0.023)	(0.022)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	0.194***	0.433***	-0.025	-0.103***
	(0.026)	(0.034)	(0.022)	(0.021)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. We transform outcomes by standardizing them with respect to the control mean instead of the logarithmic transformation that we use throughout the paper. Standard errors are clustered at the gram panchayat (GP) level. Controls include state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

D.7   Staggered Introduction of Controls

Table A21: **The Effect of Scheduled Areas on NREGS (10 km RD)**

	(1) Total	(2) Total	(3) STs	(4) STs	(5) SCs	(6) SCs	(7) Non-SCs/STs	(8) Non-SCs/STs
Panel A: Jobcards								
Scheduled Areas	-0.006	0.001	0.201***	0.207***	-0.049	-0.046	-0.097***	-0.094***
	(0.014)	(0.014)	(0.025)	(0.025)	(0.031)	(0.031)	(0.023)	(0.024)
Control Mean (Unlogged)	652.979	652.979	259.373	259.373	92.768	92.768	300.838	300.838
# GPs	14933	14933	14933	14933	14933	14933	14933	14933
# Villages	32641	32641	32641	32641	32641	32641	32641	32641
Panel B: Households Worked								
Scheduled Areas	0.014	0.011	0.198***	0.204***	-0.016	-0.023	-0.076***	-0.081***
	(0.022)	(0.023)	(0.029)	(0.029)	(0.031)	(0.031)	(0.029)	(0.029)
Control Mean (Unlogged)	220.579	220.579	98.339	98.339	29.806	29.806	92.435	92.435
# GPs	14933	14933	14933	14933	14933	14933	14933	14933
# Villages	32641	32641	32641	32641	32641	32641	32641	32641
Panel C: Workdays								
Scheduled Areas	0.009	-0.006	0.243***	0.242***	0.014	0.000	-0.098**	-0.115**
	(0.035)	(0.036)	(0.045)	(0.046)	(0.052)	(0.053)	(0.044)	(0.045)
Control Mean (Unlogged)	9748.164	9748.164	4306.585	4306.585	1259.986	1259.986	4181.593	4181.593
# GPs	14933	14933	14933	14933	14933	14933	14933	14933
# Villages	32641	32641	32641	32641	32641	32641	32641	32641
Controls	Population	Full	Population	Full	Population	Full	Population	Full

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Population Controls include ST and SC village proportion shares in 1991 and 2001. Full controls include population controls and all baseline indices. The introduction of Census controls does not seem to change point estimates substantively. All models include state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

D.8 Mahalanobis Nearest Neighbor Matching

Table A22: **The Effect of Scheduled Areas on NREGS (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	-0.053***	0.171***	-0.064**	-0.134***
	(0.017)	(0.027)	(0.028)	(0.024)
Control Mean (Unlogged)	508.322	191.572	70.847	245.903
# GPs	14933	14933	14933	14933
Panel B: Households Worked				
Scheduled Areas	-0.033	0.134***	0.015	-0.114***
	(0.025)	(0.027)	(0.025)	(0.027)
Control Mean (Unlogged)	167.636	70.017	22.919	74.700
# GPs	14933	14933	14933	14933
Panel C: Workdays				
Scheduled Areas	-0.051	0.200***	0.004	-0.145***
	(0.044)	(0.048)	(0.048)	(0.047)
Control Mean (Unlogged)	7460.817	3078.321	976.782	3405.714
# GPs	14933	14933	14933	14933

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. We match Scheduled Areas GPs to their nearest non-Scheduled Areas GP neighbor. Data are collapsed at the GP level. Nearest neighbor matching is done on all baseline indices, ST and SC village proportion, distance to border, state, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x+y+x^2+y^2+xy+x^3+y^3+x^2y+xy^2$.

D.9 Grid-level Analysis

As a robustness exercise, we create spatial grids of 1x1 degree, as shown in the map below. Table A23 shows that results are robust to including grid fixed effects.
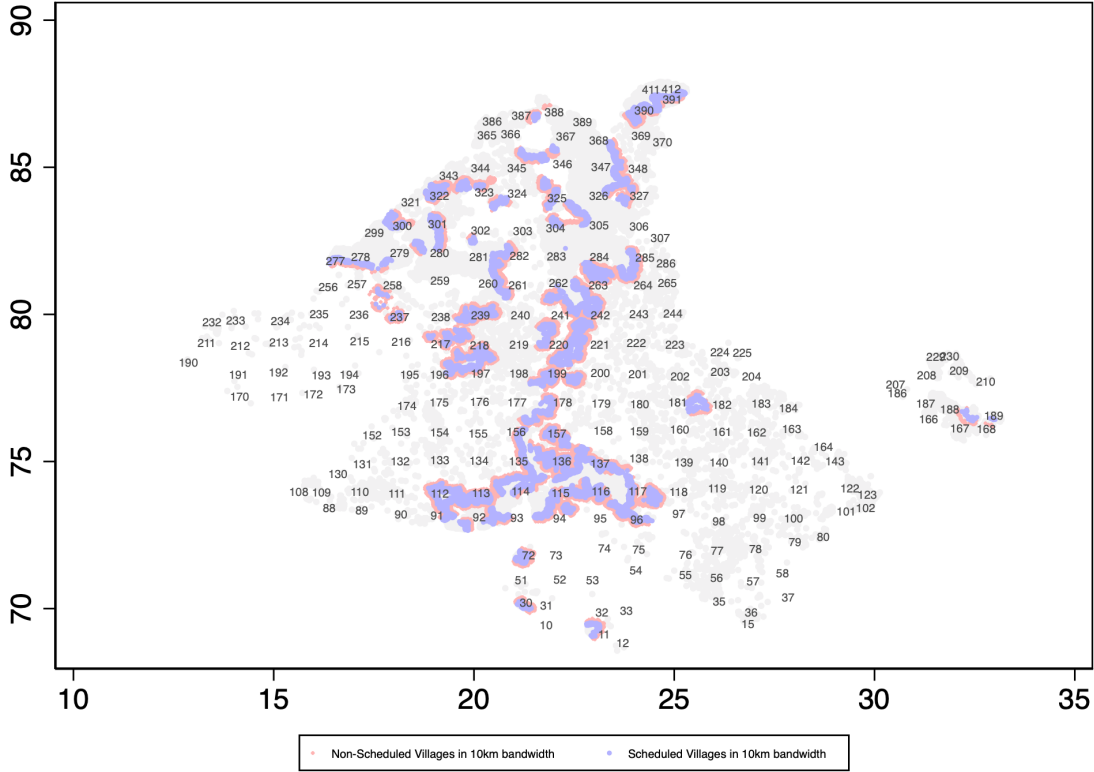

Figure A10: **10km Sample with Grids**

Notes: This figure plots the effective 10km Geographic RD sample with grid IDs. The ID labels are displayed at the centroid of sample villages in that grid.

Table A23: **The Effect of Scheduled Areas on NREGS (10 km RD) with Grid FE**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	0.008	0.220***	-0.058*	-0.066***
	(0.013)	(0.024)	(0.030)	(0.023)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	0.033	0.213***	-0.018	-0.048*
	(0.021)	(0.028)	(0.030)	(0.028)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	0.025	0.269***	-0.007	-0.077*
	(0.034)	(0.043)	(0.051)	(0.043)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, 1X1 degree grid FE, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.
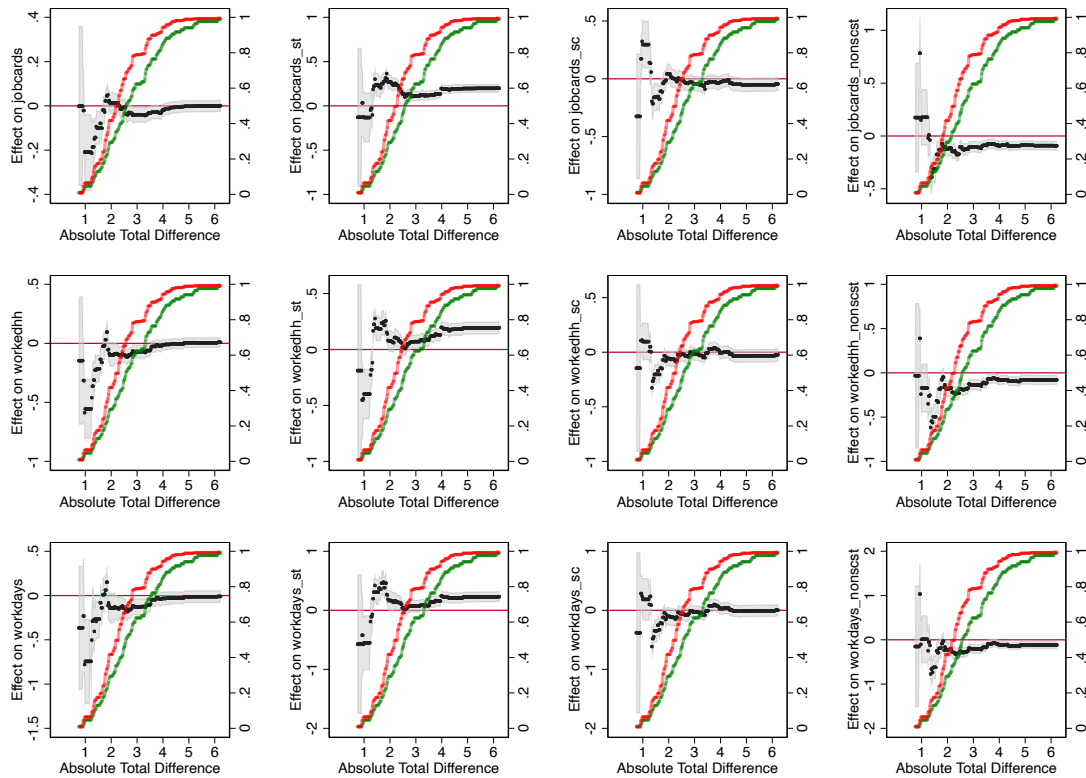
Figure A11: **Geographic RD Balance**

Notes: This figure probes the robustness of main treatment effects when the sample is restricted to various degrees of imbalance in pre-treatment Census 2001 indices. We proceed as follows: first we calculate the total absolute imbalance across all 14 Census 2001 indices. Then for each 1X1 degree grid, we calculate the degree of imbalance on this global balance index by regressing it on our standard geographic RD regression. Then we rank the districts starting at the least imbalanced on the left of the figures above to most imbalanced as we move towards the right on the x-axis. Treatment effects (on the left y-axis) are shown by the black dots for each sample while the gray area plots the associated 95 percent confidence interval. The amount of data in each sample is shown on the right y-axis. The green line traces the cumulative distribution of grids in the sample as more and more grids are added to the analysis. The red line traces the cumulative density of GPs in the analysis. Overall, we see that our results are robust and stabilize even with about 50 percent of the grids in the analysis.

Table A24: **Balance Table - 10 km RD with Grid FE & Restricted Sample**

	Difference	p-value	GPs	Villages
Population Index (Census, 2001)	0.004	0.834	9707	15930
Minority Index (Census, 2001)	0.068	0.000	9707	15930
Vulnerable Pop Index (Census, 2001)	0.046	0.000	9707	15930
Education Index (Census, 2001)	0.036	0.043	9707	15930
Medical Facilities Index (Census, 2001)	0.013	0.597	9707	15930
Water Index (Census, 2001)	0.190	0.000	9707	15930
Communications Index (Census, 2001)	-0.085	0.000	9707	15930
Banking Index (Census, 2001)	-0.056	0.004	9707	15930
Road Index (Census, 2001)	0.079	0.000	9707	15930
Urbanization Index (Census, 2001)	0.227	0.000	9707	15930
Irrigation Index (Census, 2001)	0.025	0.058	9707	15930
Agricultural Worker Index (Census, 2001)	-0.001	0.979	9707	15930
Marginal Worker Index (Census, 2001)	0.003	0.862	9707	15930
Non-Agricultural Worker Index (Census, 2001)	0.011	0.464	9707	15930
Share SC (Census, 1991)	-0.009	0.000	9707	15930
Share SC (Census, 2001)	-0.009	0.000	9707	15930
Share ST (Census, 1991)	0.083	0.000	9707	15930
Share ST (Census, 2001)	0.084	0.000	9707	15930

Notes: This table presents balance between treated and untreated units using our Geo RD specification and restricted to data from the 50 percent of the most balanced grids in Figure A11. Standard errors are clustered at the gram panchayat (GP) level. Controls include grid fixed effects. The 'Difference' column presents the treatment effect of Scheduled Areas on each Index in rows. We find that balance improves in this smaller sample. Compared to the overall balance presented in Table A12, we can see that now the difference is higher than 0.1 standard deviations for only 2 indices: water, for which the difference is smaller than for the full sample balance, and urbanization, for which the difference is slightly larger. Overall, together with the robustness of the staggered introduction of controls analysis in Appendix D.7 and several of the new robustness exercises presented in this Section, we conclude that omitted variables are unlikely to drive our results.
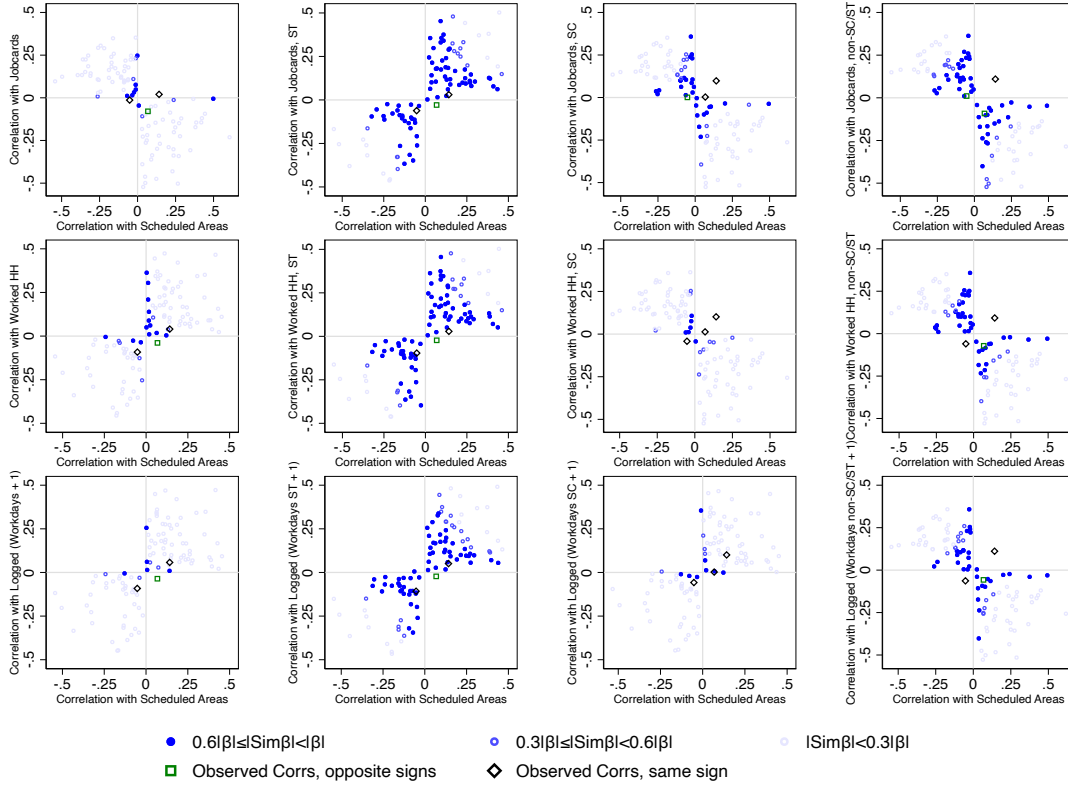
D.10 Sensitivity Analysis

Figure A12: **Sensitivity Analysis (NREGS)**

Notes: To further probe the sensitivity of our estimates, we adopt the approach suggested in Imbens (2003) to examine the degree to which an omitted confounder would impact estimated treatment effects. We also benchmark water, bank, and urban indices observed in the census against these confounders. First, we generate 200 confounders with varying degrees of correlations with the treatment and outcome variables. Second, we add these confounders in our geographic RD specification with population controls and observe what happens to the treatment effect vis-a-vis the true treatment effect. Each circle in the figure plots how correlated each confounder is with the treatment and outcomes. The color of the confounders shows the size of the simulated coefficient ($sim\beta$) with respect to the coefficient from our main specifications (β). To ease interpretation, of the four quadrants in each figure, we only plot confounders that bias the treatment effect towards zero. For instance, any confounders that are positively correlated with Workdays, ST, but negatively correlated with Scheduled Areas (or vice versa), would only increase the positive observed treatment effect of Scheduled Areas on Workdays, ST. We can see that for relatively high correlations with the treatment and outcome, the results on ST and non-SC/ST are particularly robust. To benchmark how large these correlations are, we also plot the actual correlations of the three most imbalanced census indices (water, bank, and urban) on the chart with empty squares and diamonds. If the observed correlations are in the empty quadrants, confounding from these imbalanced indices would only increase estimated treatment effects. For observed indices that lie in non-empty quadrants, we can see that for the outcomes on ST and non-SC/ST, these indices lie within the range where the estimated treatment effects shrink no more than to 60 percent of their value in the presence of confounders of this nature. Overall, we conclude that the magnitude of correlations for confounders would need to be much larger than those observed for the three most imbalanced indices for omitted variables to be an important source of bias in treatment estimates.
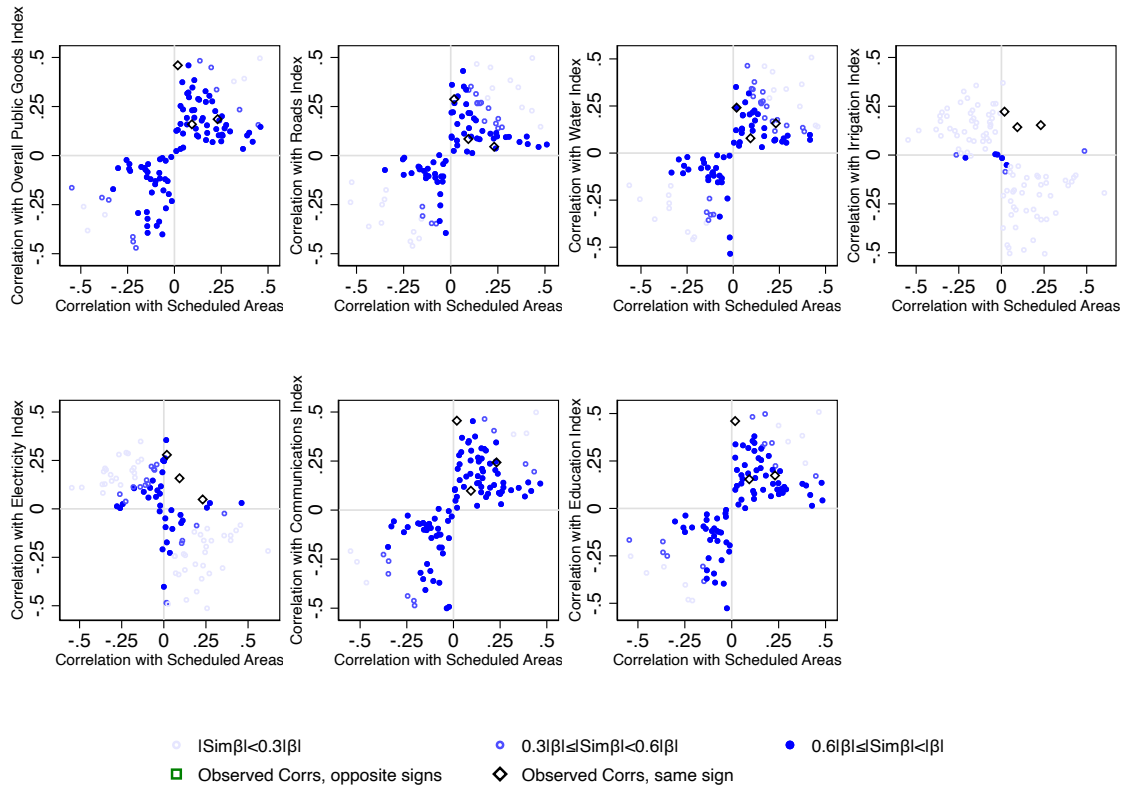
Figure A13: **Sensitivity Analysis (Public Goods, 2011)**

Notes: To further probe the sensitivity of our estimates, we adopt the approach suggested in Imbens (2003) to examine the degree to which an omitted confounder would impact estimated treatment effects. We also benchmark water, bank, and urban indices observed in the census against these confounders. First, we generate 200 confounders with varying degrees of correlations with the treatment and outcome variables. Second, we add these confounders in our geographic RD specification with population controls and observe what happens to the treatment effect vis-a-vis the true treatment effect. Each circle in the figure plots how correlated each confounder is with the treatment and outcomes. The color of the confounders shows the size of the simulated coefficient $(sim\beta)$ with respect to the coefficient from our main specifications (β). To ease interpretation, of the four quadrants in each figure, we only plot confounders that bias the treatment effect towards zero. We can see that for relatively high correlations with the treatment and outcome, the results are particularly robust. To benchmark how large these correlations are, we also plot the actual correlations of the three most imbalanced census indices (water, bank, and urban) on the chart with empty squares and diamonds. If the observed correlations are in the empty quadrants, confounding from these imbalanced indices would only increase estimated treatment effects. For observed indices that lie in non-empty quadrants, we can see that indices generally lie within the range where the estimated treatment effects shrink no more than to 60 percent of their value in the presence of confounders of this nature. Overall, we conclude that the magnitude of correlations for confounders would need to be much larger than those observed for the three most imbalanced indices for omitted variables to be an important source of bias in treatment estimates.

D.11 NREGS Results: Market Village (Census 2011 Public Goods) Sample Only

For our analysis of public goods from the 2011 Census,[40] we were only able to procure data on a smaller sample of "market villages". InfoMap India, from which our data is licensed, defines a market village as a village that holds a daily, weekly, or monthly market according to the 2011 Census. These villages are different from non-market villages in our sample in ways that one might expect: they are more populous and more developed (see Appendix table A8). Balance also appears to be better (see Appendix table A13) in market villages than for the full sample, perhaps due to the relatively specific selection criteria for market villages, although the imbalances we find are of similar magnitude – indicating that differences in balance may largely be a function of sample size.

While important in their own right, market villages are not representative of a large number of villages in India, nor, due to their comparatively higher level of development, might we expect them to be *most* in need of NREGS (though they are likely still in need). Were we to limit our analysis to market villages, the sample size would also be significantly reduced. For these reasons, we use the full sample for our analysis in the main text.

However, given the improved balance in market villages, one might wonder whether results are robust to limiting the main analysis to the market villages sample only. Indeed, Appendix Table A25 indicates that results are similar to those reported in our full sample, albeit with a stronger positive effect for STs at the extensive margin, and a lesser negative substitution effect for non-SCs/STs.

---

[40]Note that the employment results in Appendix Section E are based on a larger sample that was available for employment outcomes.

Table A25: **The Effect of Scheduled Areas on NREGS (10 km, Market Village Sub-Sample)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Log Jobcards				
Scheduled Areas	-0.030	0.275***	-0.020	-0.077
	(0.028)	(0.060)	(0.066)	(0.048)
Control Mean (Unlogged)	772.710	250.030	111.558	411.121
# GPs	2293	2293	2293	2293
# Villages	2749	2749	2749	2749
Panel B: Log Households Worked				
Scheduled Areas	0.050	0.313***	-0.016	-0.034
	(0.045)	(0.064)	(0.064)	(0.056)
Control Mean (Unlogged)	232.703	84.987	33.404	114.312
# GPs	2293	2293	2293	2293
# Villages	2749	2749	2749	2749
Panel C: Log Workdays				
Scheduled Areas	0.137*	0.494***	0.114	0.025
	(0.072)	(0.103)	(0.116)	(0.092)
Control Mean (Unlogged)	9982.793	3629.147	1368.404	4985.243
# GPs	2293	2293	2293	2293
# Villages	2749	2749	2749	2749

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors clustered by GP.

D.12 PMGSY Effects: RD Sample Only

Figure A14: **The Effect of the introduction of PESA election on PMGSY Roads (RD Sample)**
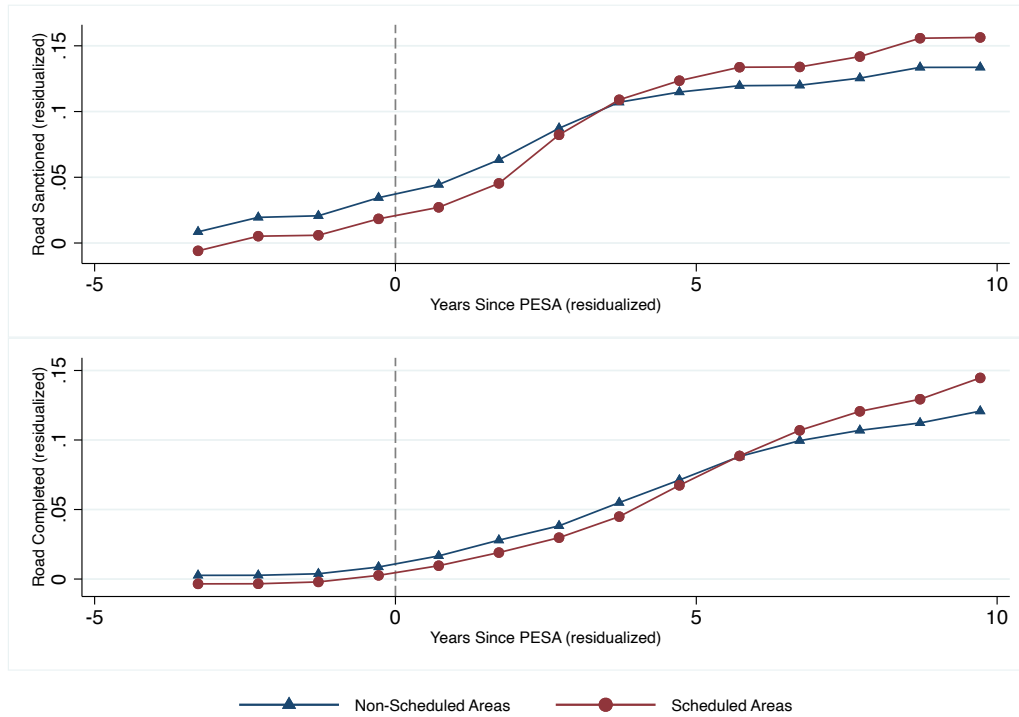
Notes: This figure plots binned means of completed PMGSY roads by Scheduled Area status on a dataset of villages in our RD sample that is residualized for village fixed effects.

D.13 District-level Muslim Rural Population Controls

Table A26: **Replication of Table 2 Main Effects on Jobcards (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Rural Muslim Population Controls				
Scheduled Areas	-0.010	0.203***	-0.042	-0.116***
	(0.014)	(0.025)	(0.031)	(0.024)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Including Squared Population Term				
Scheduled Areas	-0.010	0.202***	-0.040	-0.116***
	(0.014)	(0.025)	(0.031)	(0.024)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Controlling for Population Terciles				
Scheduled Areas	-0.000	0.206***	-0.043	-0.096***
	(0.014)	(0.025)	(0.031)	(0.023)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. This table shows that results are robust to controlling for rural Muslim district-level population in different forms.

Table A27: **Replication of Table 2 Main Effects on Worked HH (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Rural Muslim Population Controls				
Scheduled Areas	-0.019	0.180***	-0.025	-0.108***
	(0.023)	(0.029)	(0.031)	(0.030)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Including Squared Population Term				
Scheduled Areas	-0.024	0.175***	-0.026	-0.112***
	(0.023)	(0.029)	(0.031)	(0.029)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Controlling for Population Terciles				
Scheduled Areas	0.007	0.202***	-0.022	-0.085***
	(0.023)	(0.029)	(0.031)	(0.029)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. This table shows that results are robust to controlling for rural Muslim district-level population in different forms.

Table A28: **Replication of Table 2 Main Effects on Workdays (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Rural Muslim Population Controls				
Scheduled Areas	-0.053	0.200***	-0.012	-0.158***
	(0.036)	(0.046)	(0.053)	(0.046)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Including Squared Population Term				
Scheduled Areas	-0.060*	0.193***	-0.014	-0.164***
	(0.036)	(0.045)	(0.053)	(0.046)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Controlling for Population Terciles				
Scheduled Areas	-0.011	0.239***	0.002	-0.120***
	(0.036)	(0.046)	(0.053)	(0.045)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. This table shows that results are robust to controlling for rural Muslim district-level population in different forms.

E Intersecting Identities

We also investigate whether reservations have differential effects, for women and men, for several reasons. First, NREGS mandates that one-third of workers be women, and that women and men be paid equal wages. Dutta et al. (2014) find that 48% of NREGS workers are women, which is approximately twice the share of women in other casual wage work. If quotas improve program implementation, then positive effects may be particularly strong for women.

Second, minority politicians elected under quotas may be more or less responsive to women. Cassan and Vandewalle (2017) report that high caste women are less politically active than low caste women, and therefore reservations for women result in more lower caste women elected to office. Flipping this argument in our case may suggest that reservations will encourage greater participation among women.

Alternatively, men may do better where there are ST reservations. If ST are particularly in need of NREGS work, and bureaucrats are more likely to provide work for men than women, then gains in NREGS, in Scheduled Areas, may be concentrated among men. Dutta et al. (2014) report that this type of rationing is pervasive with NREGS work in poorer states.

E.1 Intersecting Identities: Decomposing Gender Effects

Do marginalized women comparatively benefit from Scheduled Areas? While "one aim of [NREGS] was to encourage women from poor households to under take work" (Jenkins and Manor, 2017, p. 174), checking this for NREGS is difficult as the data do not decompose outcomes by both identity *and* gender (see Appendix Section E.1.2).

E.1.1 Employment and Gender

With the caveat that many factors underlie changes in market conditions, we make some progress by analyzing the effects of Scheduled Areas on employment prospects by gender and types of workers in the 2011 Census. These data provide employment statistics across two categories defined by the Census: "main workers," who were employed more than 183 days, or about 6 months, in the 12 months preceding the Census, and "marginal workers," who were employed for less than 183 days.

If a large portion of individuals are solely employed through NREGS, we should expect primary gains among marginal workers due to the 100 NREGS workday maximum per household. However, individuals might also supplement their NREGS work which would make it reasonable to expect effects among main workers. Indeed, prior work shows that NREGS has positive effects on private sector employment by raising the rural reservation wage (Muralidharan, Niehaus and Sukhtankar, 2017). In addition, 70 percent of NREGS work occurs during the lean season, additively bringing new labor into the market (Jenkins and Manor, 2017, p. 170).

We observe three results in Appendix Table A29. First, consistent with the extensive margin results on NREGS, there is no effect on average employment. Second, women experience about 2.4 percent gains in employment, while men are worse off by 2.1 percent. Third, relative to the other gender, the primarily beneficiaries of Scheduled Areas are 'marginal' women workers, whose employment increases by 3%, while the primary losers are 'main' men workers.[41]

How do we interpret these results? Control means show that women are more likely to be employed as marginal workers than are men, suggesting that they work fewer days of the year on average. The treatments effects indicate that it may be these types of underemployed workers who benefit the most in Scheduled Areas, suggesting the possibility that ST women

---

[41]Appendix Figure A15 probes robustness to different bandwidths and functional forms.

Table A29: **Effects on Employment (10 km RD, Census 2011)**

	(1) Total	(2) Women	(3) Men
Panel A: Log # Overall Workers			
Scheduled Areas	-0.009	0.024*	-0.021**
	(0.010)	(0.013)	(0.010)
Control Mean (Unlogged)	552.0	231.9	319.9
# GPs	14864	14864	14864
# Villages	32522	32522	32522
H_0: $\gamma(2)$ - $\gamma(3) = 0$			$p < 0.000$
Panel B: Log # Main Workers ($>$ 183 days)			
Scheduled Areas	-0.020	0.001	-0.023
	(0.018)	(0.020)	(0.017)
Control Mean (Unlogged)	377.6	128.0	249.5
# GPs	14864	14864	14864
# Villages	32522	32522	32522
H_0: $\gamma(2)$ - $\gamma(3) = 0$			$p < 0.100$
Panel C: Log # Marginal Workers ($<$ 183 days)			
Scheduled Areas	0.001	0.030	-0.011
	(0.022)	(0.022)	(0.021)
Control Mean (Unlogged)	173.9	103.9	70.0
# GPs	14864	14864	14864
# Villages	32522	32522	32522
H_0: $\gamma(2)$ - $\gamma(3) = 0$			$p < 0.005$

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors clustered by GP. Additional controls include outcome baseline measures from the 2001 Census.

benefit more from the increase in average ST workdays.
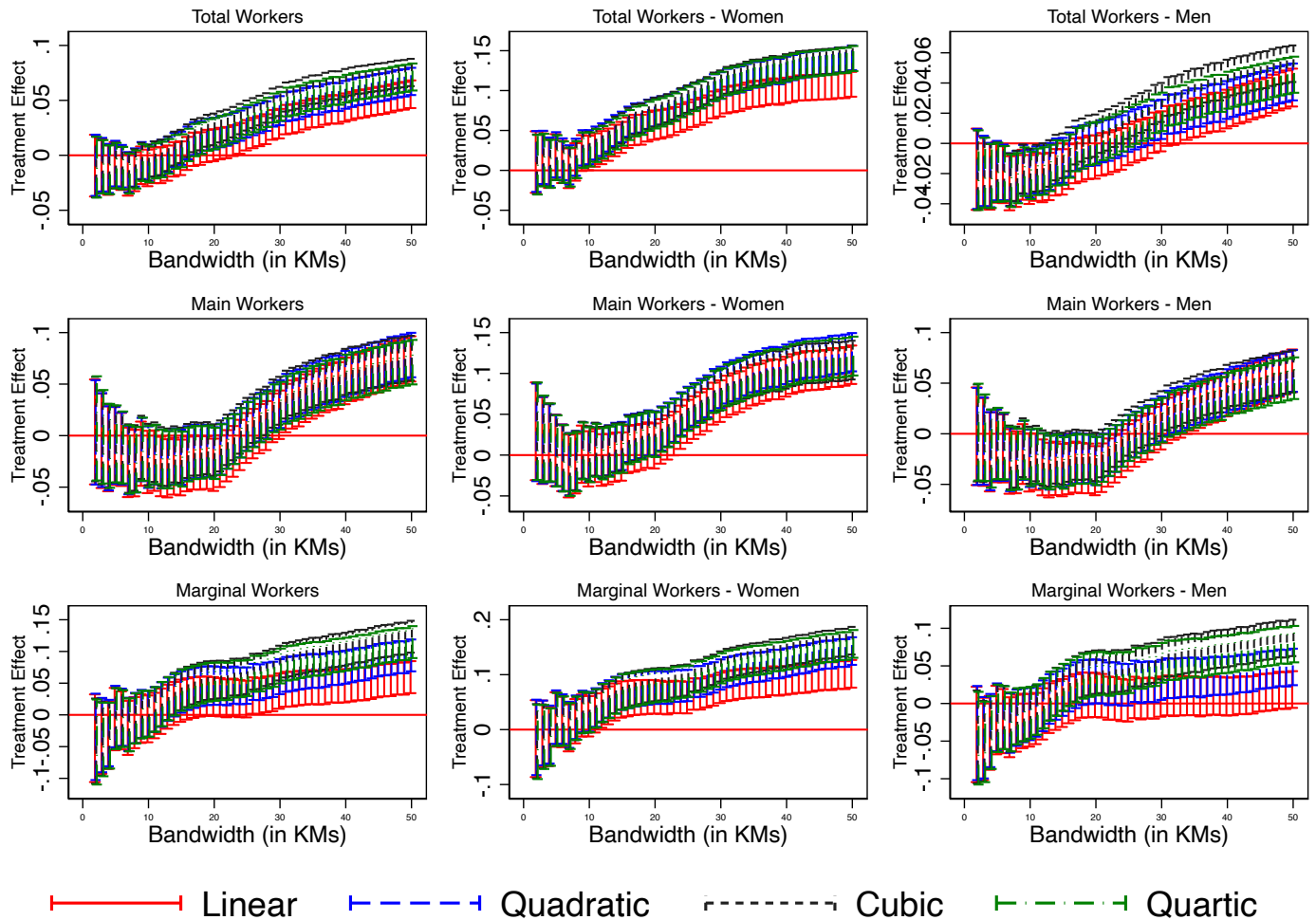
Figure A15: **The Effect of Scheduled Areas on Employment (Census 2011).** This figure plots results from a Geographic RD model with various control functions in latitudes and longitudes, as well as bandwidth. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion and state fixed effects. Main workers are those employed more than 183 years in the previous year, while Marginal Workers are those employed less than 183 days.

E.1.2 Gender NREGS Results

Appendix Table A30 presents results by gender for workdays under NREGS (the only outcome for which gender decomposed data are available) and shows that there are no key differences at least at the extensive margin across gender.

Table A30: **The Effect of Scheduled Areas on NREGS by Gender (10 km RD)**

	(1) Total	(2) Women	(3) Men
Scheduled Areas	-0.006	0.024	0.011
	(0.036)	(0.037)	(0.034)
Control Mean (Unlogged)	9748.164	4032.491	5715.673
# GPs	14933	14933	14933
# Villages	32641	32641	32641
H_0: $\gamma(2)$ - $\gamma(3) = 0$			$p = 0.37$

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.
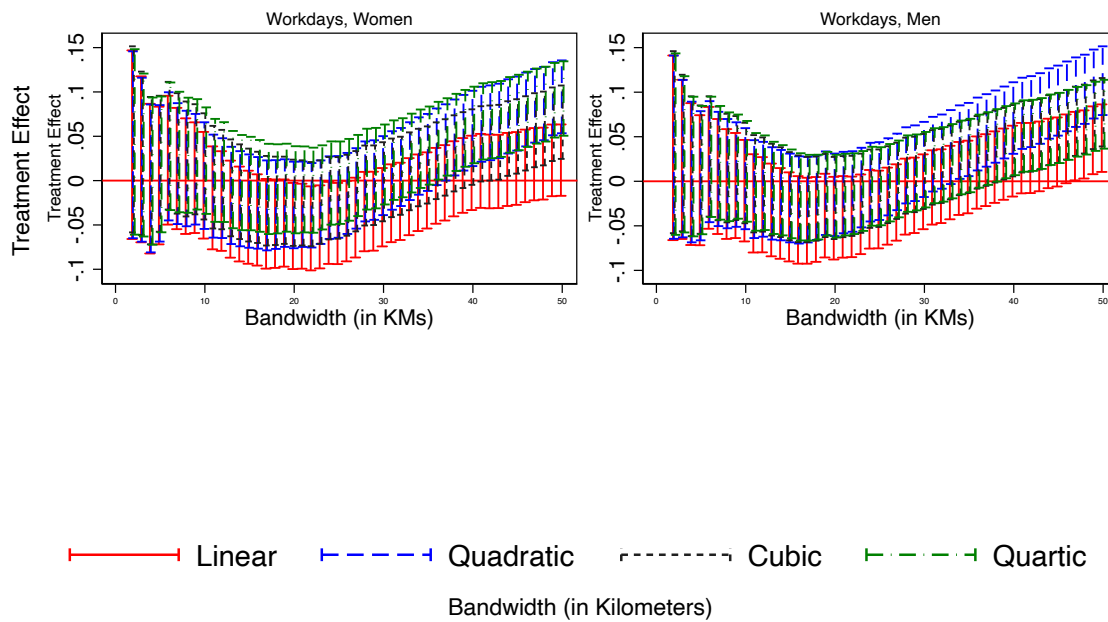
Figure A16: **RD Results by Bandwidth**

Notes: We plot the effect of Scheduled Areas on workdays for women and men at different kilometer bandwidths and with 90% confidence intervals.

F   Additional Results and Robustness of Electoral Mechanism

F.1   PESA Elections

Table A31: **The Effect of Scheduled Areas on NREGS by Number of PESA Elections (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Sch Areas × 1 PESA Election	0.058***	0.497***	-0.291***	-0.016
	(0.022)	(0.076)	(0.090)	(0.051)
Sch Areas × 2 PESA Elections	-0.022	0.287***	-0.287***	-0.462***
	(0.024)	(0.064)	(0.072)	(0.078)
Sch Areas × 3 PESA Elections	-0.007	0.139***	0.033	-0.062**
	(0.017)	(0.028)	(0.036)	(0.027)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
Panel B: Worked HH				
Sch Areas × 1 PESA Election	-0.032	0.520***	-0.474***	-0.188***
	(0.052)	(0.101)	(0.099)	(0.070)
Sch Areas × 2 PESA Elections	0.026	0.347***	-0.251***	-0.404***
	(0.039)	(0.063)	(0.081)	(0.086)
Sch Areas × 3 PESA Elections	0.017	0.123***	0.095***	-0.018
	(0.027)	(0.032)	(0.034)	(0.034)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
Panel C: Workdays				
Sch Areas × 1 PESA Election	-0.119*	0.504***	-0.619***	-0.225**
	(0.071)	(0.148)	(0.165)	(0.091)
Sch Areas × 2 PESA Elections	0.114*	0.471***	-0.383**	-0.483***
	(0.059)	(0.103)	(0.153)	(0.141)
Sch Areas × 3 PESA Elections	0.001	0.161***	0.172***	-0.045
	(0.044)	(0.051)	(0.057)	(0.053)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

F.2 ST Plurality Information

Table A32 shows the broad coverage of the data in terms of ST plurality status and being declared a Scheduled Area.

Table A32: **Number of Villages (and Gram Panchayats) and ST Share**

Panel A: Number of Observations

	Scheduled Areas	Non-Scheduled Areas
ST Plurality	8,686 (3,952)	8,954 (4,303)
ST Non-Plurality	5,580 (3,495)	9,421 (5,669)

Panel B: Average ST Share

	Scheduled Areas	Non-Scheduled Areas
ST Plurality	0.85	0.83
ST Non-Plurality	0.16	0.13

Table A33: **Treatment Effects by ST Majority (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	0.005	0.032	-0.059	-0.074**
	(0.018)	(0.031)	(0.045)	(0.036)
ST Non-Majority	0.007	0.235***	-0.052	-0.105***
	(0.021)	(0.039)	(0.043)	(0.034)
Scheduled × ST Non-Majority	-0.009	0.378***	0.028	-0.044
	(0.022)	(0.044)	(0.049)	(0.040)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
Panel B: Worked HH				
Scheduled Areas	0.006	0.057	-0.002	-0.060
	(0.028)	(0.037)	(0.044)	(0.040)
ST Non-Majority	-0.062*	0.115**	-0.061	-0.176***
	(0.037)	(0.045)	(0.042)	(0.043)
Scheduled × ST Non-Majority	0.010	0.319***	-0.047	-0.044
	(0.037)	(0.048)	(0.047)	(0.046)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
Panel C: Workdays				
Scheduled Areas	0.011	0.049	0.052	-0.074
	(0.044)	(0.054)	(0.073)	(0.061)
ST Non-Majority	-0.088	0.263***	-0.126*	-0.279***
	(0.061)	(0.073)	(0.074)	(0.069)
Scheduled × ST Non-Majority	-0.035	0.419***	-0.110	-0.087
	(0.059)	(0.077)	(0.082)	(0.072)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. ST Non-Majority defined as less than 50% of the village population share.

F.3 AC-level Reservations

Table A34: **Controlling for ST Quota in Assembly Constituency (10 km RD)**

	(1) Total	(2) ST	(3) SC	(4) Non-SC/ST
Panel A: Job Cards				
Scheduled Areas	-0.020	0.136***	-0.041	-0.083***
	(0.014)	(0.025)	(0.032)	(0.024)
AC Reserved, ST	0.197***	0.681***	-0.054	-0.107***
	(0.014)	(0.027)	(0.035)	(0.026)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	-0.033	0.128***	-0.031	-0.088***
	(0.023)	(0.030)	(0.032)	(0.030)
AC Reserved, ST	0.417***	0.725***	0.077**	0.070**
	(0.025)	(0.032)	(0.035)	(0.035)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	-0.067*	0.130***	-0.012	-0.129***
	(0.036)	(0.046)	(0.053)	(0.045)
AC Reserved, ST	0.585***	1.082***	0.122**	0.138***
	(0.038)	(0.048)	(0.058)	(0.050)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

Table A35: **Assembly Constituency SC Reservation Analysis (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	-0.008	0.182***	-0.027	-0.105***
	(0.014)	(0.026)	(0.032)	(0.024)
AC Reserved, SC	-0.106***	-0.303***	0.130**	-0.117***
	(0.026)	(0.066)	(0.064)	(0.037)
Scheduled X AC Reserved, SC	0.143***	0.349***	-0.469***	0.185
	(0.053)	(0.119)	(0.154)	(0.116)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	0.000	0.191***	-0.004	-0.094***
	(0.023)	(0.030)	(0.032)	(0.030)
AC Reserved, SC	-0.151***	-0.234***	0.160**	-0.151**
	(0.058)	(0.080)	(0.081)	(0.064)
Scheduled X AC Reserved, SC	0.102	0.035	-0.398**	0.204
	(0.105)	(0.144)	(0.175)	(0.165)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	-0.026	0.222***	0.020	-0.139***
	(0.036)	(0.047)	(0.054)	(0.046)
AC Reserved, SC	-0.270***	-0.387***	0.149	-0.283***
	(0.090)	(0.129)	(0.128)	(0.098)
Scheduled X AC Reserved, SC	0.240	0.041	-0.438	0.378*
	(0.150)	(0.222)	(0.283)	(0.224)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

Table A36: **Assembly Constituency ST Reservation Analysis: PMGSY (10 km RD)**

	(1) Roads Completed
Scheduled Areas	0.047***
	(0.006)
AC Reserved, ST	0.042***
	(0.006)
Scheduled X AC Reserved, ST	-0.037***
	(0.009)
Non-Scheduled Mean	0.127
# GPs	14933
# Villages	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.
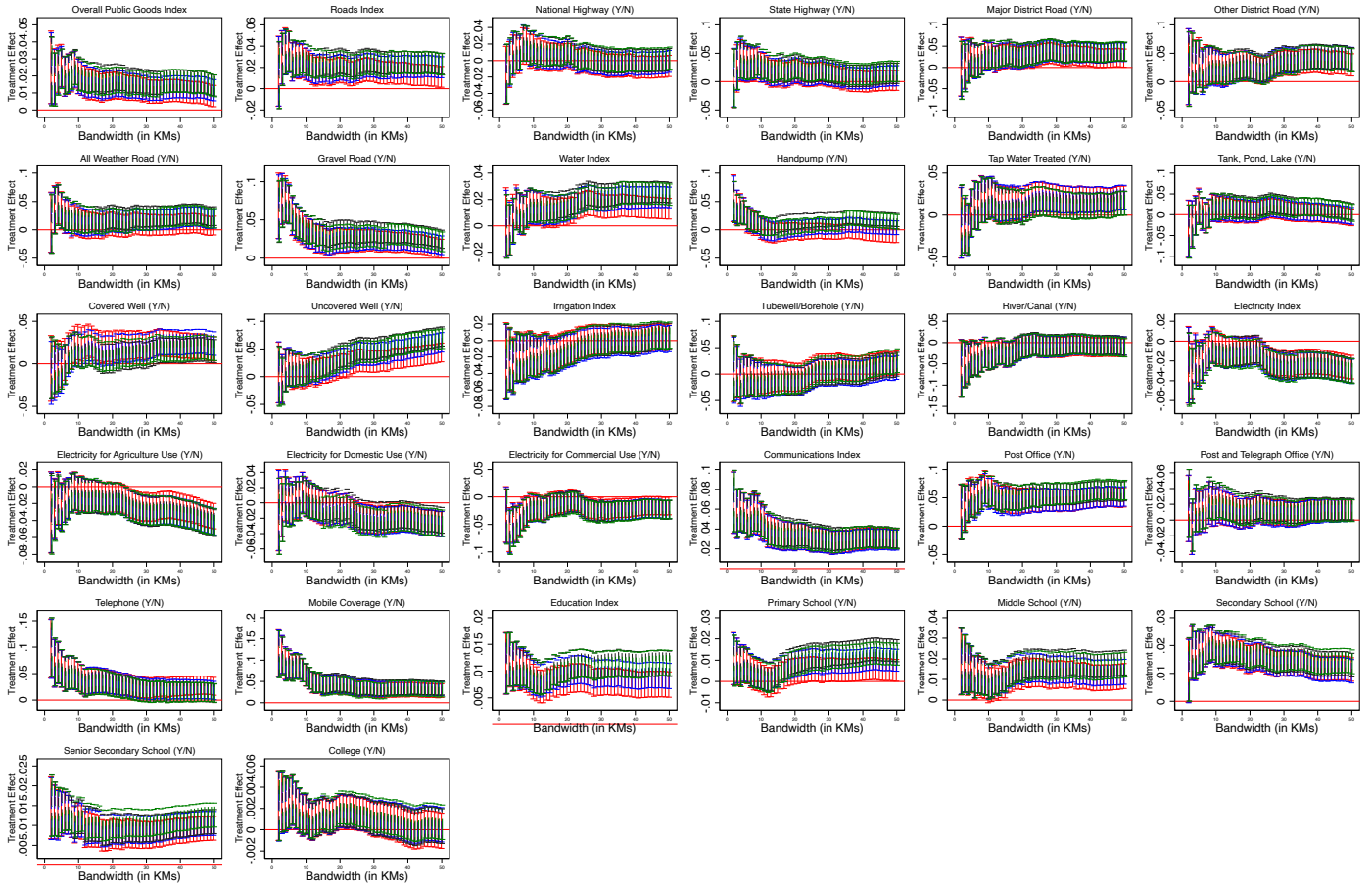
F.4 Effects on Public Goods: 2011 Census

F.4.1 Controlling for Baseline Values

Figure A17: **The Effect of Scheduled Areas on Public Goods (Census 2011).**
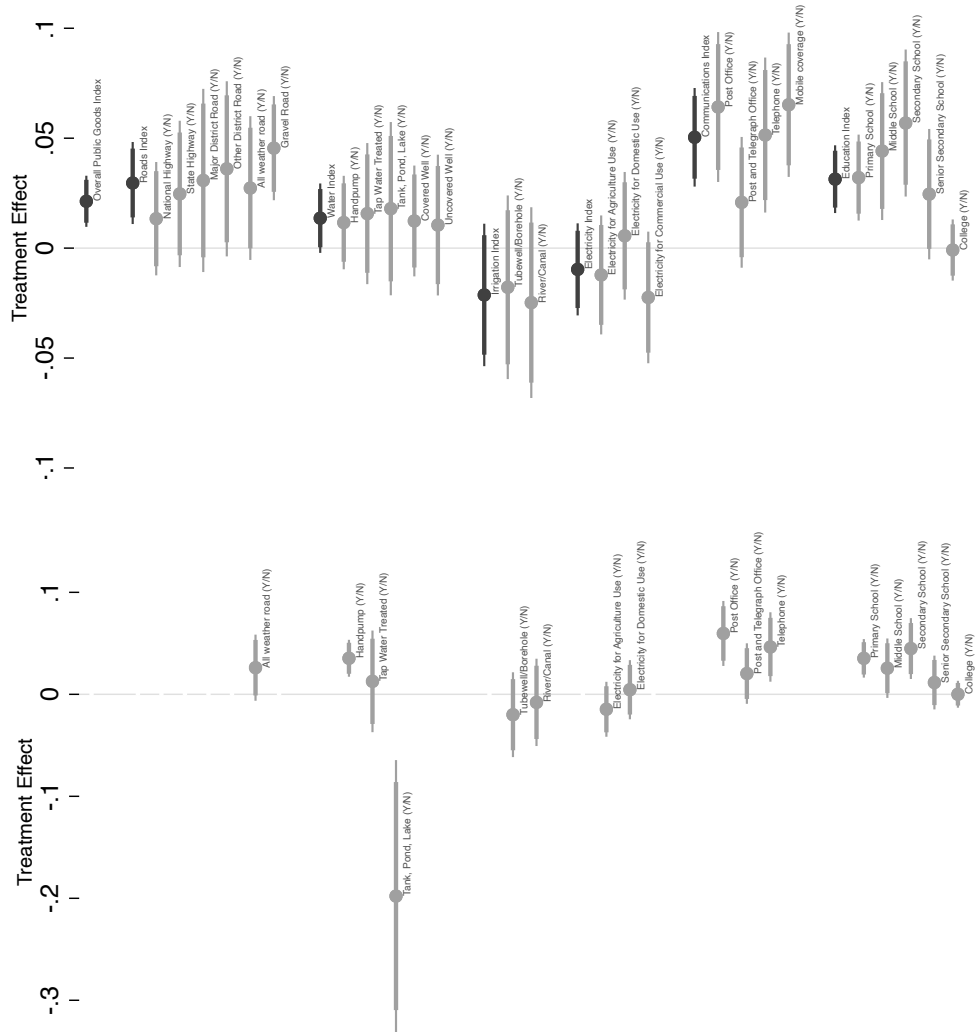
Notes: Bottom panel replicates top panel results controlling for 2001 baseline values (subject to data availability).

Figure A18: **The Effect of Scheduled Areas on Public Goods (Census 2011).** This figure plots results from a Geographic RD model with a cubic control function in latitudes and longitudes, and a 10 km bandwidth as discussed in the text. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion and state fixed effects.

F.4.2 Regression Analyses

Table A37: **Effects on Roads (10 km RD, Census 2011)**

	(1) Roads Index	(2) National Highway	(3) State Highway	(4) Major District	(5) Other District	(6) All Weather	(7) Gravel
Scheduled Areas	0.030***	0.013	0.025	0.031	0.036*	0.027	0.045***
	(0.010)	(0.013)	(0.017)	(0.021)	(0.020)	(0.017)	(0.012)
Control Mean	0.487	0.084	0.179	0.392	0.615	0.782	0.868
# GPs	2293	2293	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749	2749	2749

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

Table A38: **Effects on Water (10 km RD, Census 2011)**

	(1) Water Index	(2) Hand Pump	(3) Tap Water Treated	(4) Tank/Pond /Lake	(5) Covered Well	(6) Uncovered Well
Scheduled Areas	0.014*	0.012	0.016	0.018	0.012	0.010
	(0.008)	(0.011)	(0.016)	(0.020)	(0.013)	(0.016)
Control Mean	0.529	0.890	0.302	0.492	0.165	0.794
# GPs	2293	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749	2749

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

Table A39: **Effects on Irrigation (10 km RD, Census 2011)**

	(1) Irrigation Index	(2) Tube-well/ Borehole	(3) River/ Canal
Scheduled Areas	-0.021	-0.018	-0.025
	(0.017)	(0.021)	(0.022)
Control Mean	0.390	0.415	0.365
# GPs	2293	2293	2293
# Villages	2749	2749	2749

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

Table A40: **Effects on Electricity (10 km RD, Census 2011)**

	(1) Electricity Index	(2) Agri Use	(3) Domestic Use	(4) Commercial Use
Scheduled Areas	-0.010	-0.012	0.006	-0.022
	(0.011)	(0.014)	(0.015)	(0.015)
Control Mean	0.699	0.673	0.850	0.574
# GPs	2293	2293	2293	2293
# Villages	2749	2749	2749	2749

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

Table A41: **Effects on Communications (10 km RD, Census 2011)**

	(1) Communications Index	(2) Post- Office	(3) Post/Telegraph Office	(4) Telephone	(5) Mobile Coverage
Scheduled Areas	0.050***	0.064***	0.021	0.051***	0.065***
	(0.011)	(0.017)	(0.015)	(0.018)	(0.017)
Control Mean	0.460	0.326	0.238	0.528	0.748
# GPs	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

Table A42: **Effects on Education (10 km RD, Shrug Census 2011)**

	(1) Education Index	(2) Primary School	(3) Middle School	(4) Secondary School	(5) Senior Sec School	(6) College
Scheduled Areas	0.031***	0.032***	0.044***	0.057***	0.025	-0.001
	(0.008)	(0.010)	(0.016)	(0.017)	(0.015)	(0.007)
Control Mean	0.475	0.901	0.744	0.450	0.246	0.032
# GPs	2293	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749	2749

Notes: $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

F.4.3 Regression Analyses Robustness: Staggered Introduction of Controls

Table A43: **Effects on Public Goods with Staggered Controls (10 km RD, Shrug Census 2011)**

	(1) Overall Index	(2) Road Index	(3) Water Index	(4) Irrigation Index	(5) Electricity Index	(6) Comms Index	(7) Education Index
Panel A: No census 2001 index controls							
Scheduled Areas	0.027***	0.028***	0.021***	-0.003	-0.009	0.057***	0.040***
	(0.006)	(0.010)	(0.008)	(0.016)	(0.011)	(0.012)	(0.009)
Control Mean	0.506	0.487	0.529	0.390	0.699	0.460	0.475
# GPs	2293	2293	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749	2749	2749
Panel B: All census 2001 index controls							
Scheduled Areas	0.021***	0.030***	0.014*	-0.021	-0.010	0.050***	0.031***
	(0.006)	(0.010)	(0.008)	(0.017)	(0.011)	(0.011)	(0.008)
Control Mean	0.506	0.487	0.529	0.390	0.699	0.460	0.475
# GPs	2293	2293	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749	2749	2749

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. The index in the first column is a standardized mean of the other variables.

F.4.4 Regression Analyses Robustness: Standard Deviation Units

Table A44: **Effects on Public Goods with Standardized Outcomes (10 km RD, Shrug Census 2011)**

	(1) Overall Index	(2) Road Index	(3) Water Index	(4) Irrigation Index	(5) Electricity Index	(6) Comms Index	(7) Education Index
Scheduled Areas	0.101***	0.119***	0.064*	-0.059	-0.025	0.149***	0.119***
	(0.028)	(0.038)	(0.038)	(0.046)	(0.028)	(0.034)	(0.030)
Control Mean	-0.000	0.000	-0.000	0.000	0.000	-0.000	0.000
# GPs	2293	2293	2293	2293	2293	2293	2293
# Villages	2749	2749	2749	2749	2749	2749	2749

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2 y + xy^2$. The index in the first column is a standardized mean of the other variables. All outcomes are standardized with respect to the control group mean.

F.5 State by State Analysis

Table A45: **Effects on Logged Jobcards (State by State, 10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Sch × Andhra Pradesh	0.058	0.497***	-0.671***	-0.845***
	(0.040)	(0.123)	(0.123)	(0.157)
Sch × Chattisgarh	-0.043*	0.067	0.062	0.051
	(0.026)	(0.053)	(0.080)	(0.048)
Sch × Gujrat	0.153***	0.500***	-0.131*	-0.119*
	(0.044)	(0.066)	(0.074)	(0.068)
Sch × Himachal Pradesh	-0.645***	0.704**	-0.501	-2.810***
	(0.138)	(0.319)	(0.393)	(0.266)
Sch × Jharkhand	0.055**	0.492***	-0.277***	-0.021
	(0.022)	(0.076)	(0.090)	(0.051)
Sch × Maharashtra	-0.156***	0.022	0.026	0.036
	(0.031)	(0.058)	(0.056)	(0.048)
Sch × Madhya Pradesh	-0.058*	-0.118***	0.269***	-0.074*
	(0.031)	(0.045)	(0.057)	(0.042)
Sch × Odisha	0.082***	0.440***	-0.253**	-0.138*
	(0.031)	(0.069)	(0.100)	(0.075)
Sch × Rajasthan	0.119***	0.195	-0.044	0.144
	(0.041)	(0.130)	(0.197)	(0.128)

Notes: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Model is run without a constant. The interaction terms reflect the differences from mean control values of outcome in each state. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

Table A46: **Effects on Logged Worked HH (State by State, 10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Sch × Andhra Pradesh	0.179***	0.513***	-0.563***	-0.691***
	(0.061)	(0.116)	(0.142)	(0.167)
Sch × Chattisgarh	-0.033	0.160***	0.063	0.048
	(0.048)	(0.062)	(0.086)	(0.070)
Sch × Gujrat	0.015	0.144*	-0.125**	-0.106
	(0.096)	(0.087)	(0.054)	(0.082)
Sch × Himachal Pradesh	-0.859***	0.912**	-0.786*	-3.035***
	(0.225)	(0.361)	(0.453)	(0.282)
Sch × Jharkhand	-0.037	0.517***	-0.463***	-0.196***
	(0.052)	(0.101)	(0.099)	(0.070)
Sch × Maharashtra	-0.302***	-0.148**	-0.080*	-0.090
	(0.065)	(0.058)	(0.042)	(0.063)
Sch × Madhya Pradesh	-0.018	-0.050	0.339***	-0.085**
	(0.031)	(0.047)	(0.056)	(0.043)
Sch × Odisha	0.290***	0.598***	-0.112	0.074
	(0.052)	(0.083)	(0.106)	(0.107)
Sch × Rajasthan	0.665***	0.771***	0.641***	0.883***
	(0.078)	(0.134)	(0.187)	(0.159)

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Model is run without a constant. The interaction terms reflect the differences from mean control values of outcome in each state. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

Table A47: **Effects on Logged Workdays (State by State, 10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Sch × Andhra Pradesh	0.305***	0.723***	-1.145***	-0.955***
	(0.088)	(0.194)	(0.268)	(0.281)
Sch × Chattisgarh	0.034	0.191**	0.327**	0.180*
	(0.079)	(0.093)	(0.158)	(0.101)
Sch × Gujrat	-0.077	0.358**	-0.281**	-0.326**
	(0.159)	(0.152)	(0.129)	(0.155)
Sch × Himachal Pradesh	-0.931***	1.244**	-1.036	-3.792***
	(0.327)	(0.613)	(0.882)	(0.559)
Sch × Jharkhand	-0.127*	0.500***	-0.604***	-0.238***
	(0.071)	(0.147)	(0.165)	(0.091)
Sch × Maharashtra	-0.417***	-0.221**	-0.161*	-0.069
	(0.111)	(0.111)	(0.093)	(0.115)
Sch × Madhya Pradesh	-0.045	-0.103	0.529***	-0.164**
	(0.049)	(0.070)	(0.091)	(0.068)
Sch × Odisha	0.336***	0.657***	-0.052	0.089
	(0.070)	(0.117)	(0.155)	(0.134)
Sch × Rajasthan	1.254***	1.386***	1.547***	1.679***
	(0.102)	(0.194)	(0.349)	(0.198)

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Model is run without a constant. The interaction terms reflect the differences from mean control values of outcome in each state. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

F.6 District-level Heterogeneous Effects by ST 2001 Homicide

As noted in the text, our finding that non-SC/ST individuals received less work under NREGS could be consistent with multiple stories – which in turn could have important implications for how we interpret results. In particular, while we argue in the paper that the design of NREGS makes a demand-led opt-out mechanism unlikely (see Section), it is possible that status-quo privileged groups will reduce their demand for work under ST politicians due to associated social stigma (Akerlof and Kranton, 2010; Moffitt, 1983; Gille, 2013).

To obtain some suggestive evidence on this question, we consider whether there are heterogeneous effects based on a district's level of prejudice against STs. Should non-SC/ST individuals opting out of NREGS work be responsible for the negative substitution effect that we document, one might expect for the negative NREGS effect for non-SC/STs in Scheduled Areas to be greater in those districts that are more prejudiced against STs (and thus where there might be greater social stigma associated with obtaining work from a ST politician).

We use murders of STs as a proxy for ST prejudice (Bros and Couttenier, 2015; Sharma, 2015). To do so, we merged district-level data from the National Crime Records Bureau (NCRB) on crimes against STs and SCs at the hands of non-SCs and STs. We then used the same method as Bros and Couttenier (2015); Sharma (2015), and by the NCRB, to calculate a homicide rate per 100,000 individuals against STs: for each district, we divide the number of murders in 2001 against STs by the district-level ST population, and then we multiply the result by 100,000.[42] Compared with other outcomes, homicide data is less likely to suffer from reporting bias (Bros and Couttenier, 2015).

Homicide rates in our sample bear similarity to the descriptive statistics provided in Bros and Couttenier (2015), who used data from 2001 and 2011 across a full sample of nearly 600

---

[42]We were unable to download district-level ST and SC population data from the online portal, and we thus estimate it by aggregating up from the villages in our sample.

districts. They write, "District-wise the SC/ST homicide rates range from 0 to 8.55 per 100,000 inhabitants with an average of 0.34 and a standard deviation of 0.70" (Bros and Couttenier, 2015, 553). In our sample, district-wise SC/ST homicide rates range from 0 to 3.82 per 100,000 inhabitants with an average of 0.44 and standard deviation of 0.67. To simplify our analysis, we define a district as either having a homicide against a ST member (=1) or not (=0). Approximately 1/3 of our district had experienced a homicide against a ST.

We report results in Appendix Table A48 below. We do not find that the ST homicide rate interacts with Scheduled Areas status to produce different NREGS distributional outcomes for ST, SC, or non-SC/ST. We do find some evidence that Scheduled Areas with higher ST homicide rates perform better at the extensive margin. Our evidence thus appears to be consistent with either ST politicians changing the *supply* of NREGS work and/or with ST individuals demanding more work from village representatives who belong to their caste (Gille, 2018). Our results do *not* appear to be consistent with members of non-SCs/STs opting out of NREGS work due to the stigma attached from demanding work from a ST politician.

Table A48: **Treatment Effects by ST Homicide Rate (10 km RD)**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	-0.021	0.244***	-0.025	-0.101***
	(0.016)	(0.031)	(0.035)	(0.027)
Any ST Homicide	-0.074***	0.001	-0.290***	-0.034
	(0.022)	(0.041)	(0.052)	(0.038)
Scheduled × Any ST Homicide	0.066**	-0.110**	-0.050	0.021
	(0.030)	(0.054)	(0.069)	(0.053)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
Panel B: Worked HH				
Scheduled Areas	0.004	0.233***	-0.015	-0.070*
	(0.028)	(0.036)	(0.037)	(0.036)
Any ST Homicide	-0.090**	0.089**	-0.249***	-0.091**
	(0.036)	(0.044)	(0.049)	(0.045)
Scheduled × Any ST Homicide	0.025	-0.090	-0.014	-0.029
	(0.047)	(0.061)	(0.067)	(0.061)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
Panel C: Workdays				
Scheduled Areas	-0.010	0.290***	-0.007	-0.088*
	(0.043)	(0.057)	(0.062)	(0.054)
Any ST Homicide	-0.141**	0.060	-0.458***	-0.140**
	(0.058)	(0.069)	(0.084)	(0.071)
Scheduled × Any ST Homicide	0.019	-0.141	0.042	-0.071
	(0.073)	(0.094)	(0.113)	(0.095)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors clustered by GP.

G   Alternative Explanations

G.1   Data Manipulation and Collusion

One concern with the results is that politicians can steal funds earmarked for NREGS by convincing bureaucrats to report workdays for "ghost workers" who only exist on paper Bhatia and Dreze (2006). Importantly, the concern for the present paper is that this tendency might exist differentially more for Scheduled Areas. One validated way to measure manipulation of official data is to consider the distribution of digits in that data. The idea is that the distribution in naturally-occurring data should follow certain patterns, such as the well known Benford's law. These procedures have been adopted to study, for instance, the presence of fraud and electoral manipulation (Beber and Scacco (2012)).

For our purposes, what matters less is whether there exists fraud in NREGS data – many accounts say that there is – but, whether fraud is more or less likely to exist in Scheduled Areas. Thus, we focus not on how far the official data reported is from the "correct" distribution of numbers, but rather on whether the distribution of digits varies significantly between Scheduled and non-Scheduled Areas. This approach is particularly desirable because detecting data manipulation does not demand that we rely on further human-generated audit data.

To conduct our analysis, we consider if the deposits data in NREGS is more likely to end with a 0 in Scheduled Areas than in non-Scheduled Areas.[43] Deposits end in 0 in approximately 21% of official data in non-Scheduled Areas, but we do not find that this percentage differs across non-Scheduled and Scheduled Areas, allaying potential concerns about differential data manipulation.

As a secondary test, in Appendix Table A50, we evaluate whether the ratio of workdays to deposits varies by Scheduled Areas (column 1). We find a negligible 2 percentage point

---

[43]Previous work has shown that the distribution of manipulated data often bunches at 0 (Best et al. (2015)).

difference (relative to a control mean of 81.8 percent). In column 2, we also test if deposits data are missing differently by treatment and do not find evidence that this is the case.

Table A49: **Random Digit Test**

	(1) Amount Disbursed Ending in 0
Scheduled Areas	0.012
	(0.009)
Control Mean (Unlogged)	0.216
# GPs	14933
# Villages	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Outcome is whether the deposit data in NREGS ends in 0. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

Table A50: **Missingness in Deposits**

	(1) Workdays to Disbursements Ratio	(2) Missing Obs in Disbursments
Scheduled Areas	-0.020***	-0.003
	(0.004)	(0.004)
Control Mean	0.818	0.133
# GPs	12396	14933
# Villages	28616	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$.

G.2 Distance from the Center

Scheduled Areas are more rural and may be less reliant on centralized government than non-Scheduled Areas. Because NREGS implementation is primarily determined at the local level, differences in reliance on centralized government – rather than quotas – may be driving our observed effects. To evaluate this possibility, we control for each village's distance to the most populous area in our data and find in Table A51 that our results are unchanged.

Table A51: **The Effect of Scheduled Areas on NREGS (10 km RD), Controlling for Distance from Block Center**

	(1) Total	(2) STs	(3) SCs	(4) Non-SCs/STs
Panel A: Jobcards				
Scheduled Areas	0.001	0.201***	-0.048	-0.094***
	(0.014)	(0.025)	(0.031)	(0.024)
Control Mean (Unlogged)	652.979	259.373	92.768	300.838
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel B: Households Worked				
Scheduled Areas	0.014	0.200***	-0.021	-0.076***
	(0.023)	(0.029)	(0.031)	(0.029)
Control Mean (Unlogged)	220.579	98.339	29.806	92.435
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641
Panel C: Workdays				
Scheduled Areas	0.001	0.240***	0.006	-0.107**
	(0.036)	(0.046)	(0.053)	(0.045)
Control Mean (Unlogged)	9748.164	4306.585	1259.986	4181.593
# GPs	14933	14933	14933	14933
# Villages	32641	32641	32641	32641

Notes: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Standard errors are clustered at the gram panchayat (GP) level. Controls include all baseline indices, ST and SC village proportion, state fixed effects, and a flexible function in village centroid longitudes (x) and latitudes (y) of the form: $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. Finally, these regressions also control for each village's distance from its block's most populous village.

H Constructing Census Indices

Due to the large number of variables found in the 2001 Indian census (203 total variables across the Socio-Demographic module and the Infrastructure module), we decided to combine the individual census variables into indices. Where possible, we generated these indices using the Indian Census' variable groupings - for instance, the Education Index combines 10 variables listed in the census in a group which all refer to different measures of school and college facilities. However, we note that the titles we assign to indices – for example, "Banking Index", should be taken as suggestive. Each index is constructed according to the following steps: first, we standardize and center each constituent variable. Second, we assign missing variable values the treatment group mean if at least one index component is non-missing for that observation. Third, we standardize each variable to the control group mean and standard deviation. Fourth, we take as our index value the average of all the constituent variable values (Kling, Liebman and Katz, 2007).

We list each census' variable included in each index below. Note while the vast majority of census variables are binary or counts, additional variables provide additional qualitative information when village data was unavailable (known as 'range codes') – we omit these non-numerical variables for simplicity.

We also present analyses using 2011 Census data in Table A29, in Figures 4, A15, and A18, and in Appendix F.4. The data at our disposal for the 2011 Census did not always match exactly the data for the 2001 Census. Accordingly, below, we note which variables we had for which Censuses, as well as the cases when it was necessary to recode variables (typically, to be binary) so that we could match data across the two Censuses.

Socio-Demographic and Economic Module of 2001 Census, and 2011 Indian Census

Key:

* = In data for 2001 Census only, ** = In data for 2011 Census only.

- Population Index

    TOT_NM_HH - Total number of households

    TOT_POP - Total population

    M_POP - Male population

    F_POP - Female population

- Vulnerability Index

    TOT_L6 - Total pop below 6 years

    F_L6 - Female pop below 6 years

    TOT_ILLT - Total Illiterates

    F_ILLT - Female Illiterates

    TOT_NNW - Total Non-workers

    F_NNW - Female Non-workers

- Minority Index

    TOT_SC - Total scheduled caste

    M_SC - Male scheduled caste

    F_SC - Female scheduled caste

    TOT_ST - Total scheduled tribe

    M_ST - Male scheduled tribe

    F_ST - Female scheduled tribe

- Agricultural Worker Index

    TOT_CULT - Total Cultivators

    M_CULT - Male Cultivators

F_CULT - Female Cultivators

TOT_AGLB - Total Agricultural Labourers

M_AGLB - Male Agricultural Labourers

F_AGLB - Female Agricultural Labourers

T_MRG_CULT - Total Marginal workers as cultivators

M_MRG_CULT - Male Marginal workers as cultivators

F_MRG_CULT - Female Marginal workers as cultivators

T_MRG_AGLB - Total Marginal workers as agricultural labourers

M_MRG_AGLB - Male Marginal workers as agricultural labourers

F_MRG_AGLB - Female Marginal workers as agricultural labourers

- Non-Agricultural Worker Index

  TOT_MFHH - Total Household industry workers

  M_MFHH - Male Household industry workers

  F_MFHH - Female Household industry workers

  TOT_OTH_W - Total other workers

  M_OTH_W - Male other workers

  F_OTH_W - Female other workers

  T_MRG_HH - Total Marginal workers household industry workers

  M_MRG_HH - Male Marginal workers household industry workers

  F_MRG_HH - Female Marginal workers household industry workers

  T_MRG_OTH - Total Marginal workers as other workers

  M_MRG_OTH - Male Marginal workers as other workers

  F_MRG_OTH - Female Marginal workers as other workers

- Marginal Worker Index

    TOT_MRW - Total Marginal workers other workers

    M_MRW - Male Marginal workers other workers

    F_MRW - Female Marginal workers other workers

Infrastructure and Amenities Module of 2001 Indian Census, and Market Villages Data for 2011 Census

- Education Index*

    EDU_FAC - Educational facilities (binary)*

    P_SCH - Number of Primary School*

    M_SCH - Number of Senior Secondary School*

    S_SCH - Number of Secondary School*

    S_S_SCH - Number of Senior Secondary School*

    COLLEGE - Number of Collage*

    IND_SCH - Number of Industrial School*

    TR_SCH - Number of Training School*

    ADLT_LT_CT - Number of Adult literacy Class/Centre*

    OTH_SCH - Number of Other educational facilities*

- Medical Facilities Index*

    MEDI_FAC - Medical facilities (binary)*

    ALL_HOSP - Allopathic hospital, Maternity and Child Welfare Centre and Primary Health Centre*

    AYU_HOSP - Number of Allopathic Hospital*

UN_HOSP - Number of Unani Hospital*

HOM_HOSP - Number of Homeopathic Hospital*

ALL_DISP - Number of Allopathic Dispensary*

AYU_DISP - Number of Ayurvedic Dispensary*

UN_DISP - Number of Unani Dispensary*

HOM_DISP - Number of Homeopathic Dispensary*

MCW_CNTR - Number of Maternity and Child Welfare Centre*

M_HOME - Number of Maternity Home*

CWC - Number of Child Welfare Centre Number of Health Centre*

H_CNTR - Number of Health Centre*

PH_CNTR - Number of Primary Health Centre*

PHS_CNT - Number of Primary Health Sub Centre*

FWC_CNTR - Number of Family Welfare Centre Number of T.B. Clinic*

TB_CLN - Number of T.B. Clinic*

N_HOME - Number of Nursing Home*

RMP - Number of Registered Private Medical Practitioners*

SMP - Number of Subsidized Medical Practitioners*

CHW - Number of Community Health workers*

OTH_CNTR - Number of Other medical facilities*

- Water Index

HDP_ST - Hand Pump (HP)

SPR_ST - Spring (S)

TPL_ST - Tank/Pond/Lake

A91

TWT_ST - Tap Water (Treated)

TUBEWELL - Tubewell Water (TW)

WELL - Well Water (W) (2011 covered/uncovered well variables combined)

DRNK_WAT_F - Drinking Water facility (binary)*

RIVER - River Water(R)*

OTHER - Other drinking water sources (O)*

TANK - Tank Water (TK)*

TWU_ST - Tap Water (Untreated) **

WAT_BOM - Water Bounded Macadam**

- Communications Index

  BS_FAC - Bus services (2011 Census private/public bus services variables combined)

  NAV_WRC - Navigable water way including River, Canal etc. (2001 recoded to Y/N)

  PO_ST - Number of Post Office (2001 recoded to Y/N)

  PT_OF_ST - Number of Post and Telegraph Office (2001 recoded to Y/N)

  RS_FAC - Railways services

  TEL_ST - Number of Telephone connections (2001 recoded to Y/N)

  COMM_FAC - Communication*

  P_T_FAC - Post, Telegraph and Telephone facilities (binary)*

  TELE_OFF - Number of Telegraph Office*

  AUT_MODA - Auto/Modified Autos**

  INC_CSC - Internet Cafes / Common Service Centre (CSC)**

MOB_PH_CV - Mobile Phone Coverage**

RAIL_STAT - Railway Station**

SRF_SERV - Sea/River/Ferry Service**

SPO_ST - Sub Post Office**

VANS_ST - Vans**

- Banking Index

  AGRI_CRSO - Number of Agricultural Credit Societies (2001 recoded to Y/N)

  COOP_BANK - Number of Agricultural Credit Societies (2001 recoded to Y/N)

  ST_AU_FAC - Number of Stadium/Auditorium

  BANK_FAC - Banking facility (binary)*

  COMM_BANK - Number of Cooperative Commercial Banks*

  CRSOC_FAC - Credit Societies (Y/N)*

  C_V_HALL - Number of Cinema/Video-hall*

  NAC_SOC - Number of Non Agricultural Credit Societies*

  OTHER_SOC - Number of Other Credit Societies*

  RC_FAC - Recreational and Cultural facilities (binary)*

  SP_CL_FAC - Number of Sports Club*

  ATM_ST - ATM**

- Road Index

  ALL_WE_RD - Approach - Paved (all weather) Road (2001 recoded to Y/N)

  APP_FP - Approach - Foot Path*

  APP_MR - Approach - Mud Road*

APP_NAVCAN - Approach - Navigable Canal*

APP_NAVRIV - Approach - Navigable River*

APP_NW - Approach - Navigable water-way other than river or Canal*

GRA_KU_RD - Gravel (kuchha) Roads**

NH_ST - National Highway**

MDR_ST - Major District Road**

OTH_DR - Other District Road**

SH_ST - State Highway**

- Urbanization Index

    POW_SAU - Electricity of Agricultural use (2001 recoded to Y/N)

    POW_SDU - Electricity for Domestic use (2001 recoded to Y/N)

    PAP_MAG - Newspaper/Magazine (binary)*

    POWER_ALL - Electricity for all purposes*

    POWER_OTH - Electricity of other purposes*

    POWER_SUPL - Power supply (binary)*

    A_INCEXP - Separate figures available (Y/N). If Yes:*

    TOT_EXP - Total Expenditure*

    TOT_INC - Total Income*

    POW_SCU - Power Supply For Commercial Use**

- Irrigation Index

    RIC_ST - River/Canal (2001 canal, river variables combined, recoded to Y/N)

    TWB_ST - Tube well/borehole (2001 tubewell variables combined, recoded to Y/N)

CULT_WASTE - Culturable waste (including gauchar and groves)*

LAND_FORES - Forest Irrigated (by source)*

LAKE_IRR - Lake*

TANK_IRR - Tank*

OTH_IRR - Others [Water source]*

W_FALL - Waterfall*

WELL_WO_EL - Well (without electricity)*

WELL_W_EL - Well (with electricity)*

TOT_IRR - Total Irrigated Area*

UN_IRR - Unirrigated Area*

AREA_NA_CU - Area not available for cultivation*

Appendix References

Akerlof, George A. and Rachel E. Kranton. 2010. *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton: Princeton University Press.

Bardhan, Pranab K, Dilip Mookherjee and Monica Parra Torrado. 2010. "Impact of political reservations in West Bengal local governments on anti-poverty targeting." *Journal of Globalization and development* 1(1).

Beaman, Lori, Esther Duflo, Rohini Pande, Petia Topalova et al. 2010. Political reservation and substantive representation: Evidence from Indian village councils. In *India Policy Forum 2010/11 Volume.* Vol. 7 pp. 159–191.

Beber, Bernd and Alexandra Scacco. 2012. "What the numbers say: A digit-based test for election fraud." *Political analysis* 20(2):211–234.

Besley, Timothy, Rohini Pande and Vijayendra Rao. 2005. "Participatory democracy in action: Survey evidence from South India." *Journal of the European Economic Association* 3(2-3):648–657.

Besley, Timothy, Rohini Pande and Vijayendra Rao. 2007. "Political Economy of Panchayats in South India." *Economic and Political Weekly* 42(8):661–666.

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn and Mazhar Waseem. 2015. "Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan." *Journal of political Economy* 123(6):1311–1355.

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

Béteille, André. 1986. "The concept of tribe with special reference to India." *European Journal of Sociology* 27(2):296–318.

Bhatia, Bela and Jean Dreze. 2006. "Employment Guarantee in Jharkhand: Ground Realities." *Economic and Political Weekly* 41(29):3198–3202.

Bhavnani, Rikhil R. 2009. "Do electoral quotas work after they are withdrawn? Evidence from a natural experiment in India." *American Political Science Review* 103(1):23–35.

Bros, Catherine and Mathieu Couttenier. 2015. "Untouchability, homicides and water access." *Journal of Comparative Economics* 43(3):549 – 558.
URL: *http://www.sciencedirect.com/science/article/pii/S0147596714001243*

Calonico, Sebastian, Cattaneo Mathias D. Farrell Max H. and Rocio Titiunik. 2017. "rdrobust: Software for Regression Discontinuity Designs." *Stata Journal* 17:372–404.

Cassan, Guilhem and Lore Vandewalle. 2017. "Identities and Public Policies: Unintended Effects of Political Reservations for Women in India." (18-2017).

Chauchard, Simon. 2014. "Can descriptive representation change beliefs about a stigmatized group? Evidence from rural India." *American political Science review* 108(2):403–422.

Chaudhury, Dipanjan Roy. 2005. "Autonomous District Councils and Panchayati Raj Institutions in North-East India." *Dialogue* 7(1).

Chin, Aimee and Nishith Prakash. 2011. "The redistributive effects of political reservation for minorities: Evidence from India." *Journal of Development Economics* 96(2):265–277.

Commissioner for Scheduled Castes and Scheduled Tribes. N.d. *Twenty-Fourth Report of the Commissioner for SC and ST.* New Delhi: Government of India.

Corbridge, Stuart. 2000. "Competing Inequalities: The Scheduled Tribes and the Reservations System in India's Jharkhand." *The Journal of Asian Studies* 59(1):62–85.

Corbridge, Stuart. 2002. "The Continuing Struggle for India's Jharkhand: Democracy, Decentralisation and the Politics of Names and Numbers." *Commonwealth & Comparative Politics* 40(3):55–71.

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

Damodaran, Vinita. 2011. Colonial Constructions of the 'tribe' in India: The Case of Chotanagpur. In *Adivasis In Colonial India: Survival, Resistance and Negotiation*, ed. Biswamoy Pati. Orient BlackSwan.

Duflo, Esther and Raghabendra Chattopadhyay. 2004. "Women as policy makers: Evidence from a randomized policy experiment in India." *Econometrica* 72(5):1409–1443.

Dunning, Thad. 2010. "Do quotas promote ethnic solidarity? Field and natural experimental evidence from India." *Unpublished manuscript, Yale University* .
URL: *https: // pdfs. semanticscholar. org/ be8e/ 9951c39a6d85b7ac9fec77f96af497d035ec. pdf*

Dunning, Thad and Janhavi Nilekani. 2013. "Ethnic Quotas and Political Mobilization: Caste, Parties, and Distribution in Indian Village Councils." *American Political Science Review* 107(1):3556.

Dutta, Puja, Rinku Murgai, Martin Ravallion and Dominque van de Walle. 2014. *Right to Work: Assessing India's Emloyment Guarantee Scheme in Bihar.*

Election Commission of India. 2014. "Election Results - Full Statistical Reports.".
URL: *http: // eci. nic. in*

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

Gille, Veronique. 2013. "Stigma in positive discrimination application? Evidence from quotas in education in India." *Conference Paper, Centre d'Economie de la Sorbonne* .

Gille, Veronique. 2018. "Applying for social programs in India: Roles of local politics and caste networks in affirmative action." *Journal of Comparative Economics* 46(2):436–456.

Guha, Ramachandra. 1996. "Savaging the Civilised Verrier Elwin and the Tribal Question in Late Colonial India." *Economic & Political Weekly* 31(35/37).

Gupta, Sanjukta Das. 2011*a*. *Adivasis and the Raj: Socio-economic Transition of the Hos, 1820-1932.* Orient BlackSwan.

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

H.H., Risley. 1903. *General Report by H. H. Risley and. E. A. Gait.* Calcutta: Census of India.

Htun, Mala. 2004. "Is Gender like Ethnicity? The Political Representation of Identity Groups." *Perspectives on Politics* 2(3):439–458.

ILO. 2015. "ILO-PMGSY Rural Road Project Strengthening Capacity of Panchayati Raj Institutions: Managing Maintenance of Rural Roads.".
URL: *http://pmgsy.nic.in/RSCPR.pdf*

Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *Review of Economic Studies* 79:933–959.

Imbens, Guildo W. 2003. "Sensitivity to exogeneity assumptions in program evaluation." *American Economic Review* 93(2):126–132.

Iyer, Lakshmi and Maya Reddy. 2013. "Redrawing the Lines: Did Political Incumbents Influence Electoral Redistricting in the World's Largest Democracy?" *Harvard Business School Working Paper* (14-051).
URL: *https://www.hbs.edu/faculty/Pages/item.aspx?num=46005*

Jenkins, Rob and James Manor. 2017. *Politics and the Right to Work: India's National Rural Employment Guarantee Act.* Oxford University Press.

Jensenius, Francesca R. 2015. "Development from representation? A study of quotas for the scheduled castes in India." *American Economic Journal: Applied Economics* 7(3):196–220.

Kling, Jeffrey R., Jeffrey B. Liebman and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1):83–119.

Ministry of Drinking Water and Sanitation. 2014. "National Rural Drinking Water Programme.".
URL: *https://web.archive.org/web/20140214180541/http://indiawater.gov.in/imisreports/Reports/BasicInformation/rpt_RWS_AbstractData_S.aspx?Rep=0&RP=Y&APP=IMIS*

Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma." *American Economic Review* 73(5):1023–35.

Muralidharan, Karthik, Paul Niehaus and Sandip Sukhtankar. 2017. "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India." (23838).

Nazer, M. 2004. *A Study of Land Alienation and Indebtedness among Tribals in Tamil Nadu, Kerala and Karnataka.* Planning Commission of India.

Palaniswamy, Nethra and Nandini Krishnan. 2012. "Local Politics, Political Institutions, and Public Resource Allocation." *Economic Development and Cultural Change* 60(3):449–473.

Pande, Rohini. 2003. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India." *American Economic Review* 93(4):1132–1151.

Parthasarathy, Ramya, Vijayendra Rao and Nethra Palaniswamy. 2017. "Deliberative inequality: a text-as-data study of Tamil Nadu's village assemblies." *Working Paper* .
URL: *http://documents.worldbank.org/curated/en/582551498568606865/ Deliberative-inequality-a-text-as-data-study-of-Tamil-Nadus-village-assemblies*

Pati, Biswamoy. 2011. *Adivasis In Colonial India: Survival, Resistance and Negotiation.* Orient BlackSwan.

Raabe, Katharina, Madhushree Sekher and Eva Schiffer. 2010. "How to Overcome the Governance Challenges of Implementing NREGA Insights from Bihar Using Process-Influence Mapping." *IFPRI Discussion Paper 00963* (April).

Radhakrishna, Meena. 2011. Of Apes and Ancestors: Evolutionary Science and Colonial Ethnography. In *Adivasis In Colonial India: Survival, Resistance and Negotiation*, ed. Biswamoy Pati. Orient BlackSwan.

Rudolph, Lloyd I. and Susanne Hoeber Rudolph. 2010. "Federalism as State Formation in India: A Theory of Shared and Negotiated Sovereignty." *International Political Science Review* 31(5):553–572.

Sharma, Smriti. 2015. "Caste-based crimes and economic status: Evidence from India." *Journal of Comparative Economics* 43(1):204 – 226.
URL: *http://www.sciencedirect.com/science/article/pii/S0147596714001048*

Statoids. N.d. "Districts of India.".
URL: *http://www.statoids.com/yin.html, year = 2014*

Sundar, Nandini. 2005. "Custom and Democracy in Jharkhand." *Economic and Political Weekly* 40(41):4430–4434.

Tillin, Louise. 2013. *Remapping India: New States and Their Political Origins.* New York: Oxford University Press.

Upadhya, Carol. 2009. Law, Custom, and Adivasi Identity: Politics of Land Rights in Chotanagpur. In *Legal Grounds: Natural Resources, Identity, and the Law in Jharkhand,* ed. Nandini Sundar. Oxford University Press.