Compounding effects of spatial segregation, hard rock aquifers and pumping energy on drinking water access in rural India

Sudatta Ray[1,2] and Isabella Gupta[1]

[1]Department of Geography, National University of Singapore, Singapore

[2]Environmental Studies, Division of Social Science, Yale-NUS College, Singapore

Inequality in accessing public resources is a persisting problem among marginalized communities globally. In developing countries, the lack of access extends to even the most basic services such as reliable drinking water and electricity. Using high-resolution, interdisciplinary data across seven Indian states, we find Scheduled Tribe (ST) populations to be the most spatially segregated within villages and across hard-rock aquifers. Nearly 50% of the ST population lived in villages where they made up more than 80% of the village population, compared to less than 3% of similarly concentrated villages with other marginalized communities. We find villages with high ST concentration to also house approximately 22% fewer households accessing pumped water, and instead source drinking water from uncovered wells, which suffer from low productivity, quality, and high seasonal fluctuations. We find equally poor outcomes even among states expressly created to serve ST interests making their low political representation an unlikely factor attributing to the poor outcomes. Future policies must consider biophysical elements that affect public resource access, and identify areas of overlap with the distribution of marginalized populations. These overlaps while creating unforeseen poverty traps for marginalized communities, could also offer opportunities of targeting and compounding positive impacts.

Introduction

Safe and reliable drinking water supply is one of the greatest abiding challenges of the rural developing world. In 2022, between 8%-42% of rural population on average lacked access to even basic drinking water services across Central and South Asia, North, West and Sub-Saharan Africa [1][1]. Over 90% of all rural drinking water was based on groundwater sources in the same regions [1]. Sufficiency of water volumes however is not guaranteed by mere presence of water infrastructure, latter being the most often recorded data. Groundwater levels can fluctuate by many meters across seasons at any given location influenced by precipitation and aquifer characteristics. Climate change induced uncertainty in rainfall is likely to further increase groundwater fluctuations and future stocks [2–4]. Therefore, continuous monitoring of existing water infrastructure is needed to ascertain true coverage today, and to plan for adequate water infrastructure in the future. Groundwater based drinking water supply requires two key inputs - groundwater resource and pumping energy to treat and transport it. Yet, with a few exceptions, drinking water literature rarely includes analyses of either, largely focusing on the distribution of, and access to water infrastructure [5, 6]. Lack of access to basic public resources more broadly is a persistent problem among marginalised communities globally. Competition among diverse ethnic groups (ethnic diversity) and localized concentrations of certain marginalized groups (spatial segregation) have been identified as two key factors contributing to inequities in access to resources across different geographical contexts [7–11]. However, the role of biophysical factors is largely overlooked, although an important determinant of access to water [12].

We study the distribution of drinking water infrastructure across Indian villages by creating a novel dataset including socioeconomic and biophysical data. Such granular and interdisciplinary data allow us to make three main contributions - first, we identify the association between high concentrations of marginalized populations

[1]Basic services includes an improved source and a roundtrip collection time of less than 30 minutes including queuing [1]

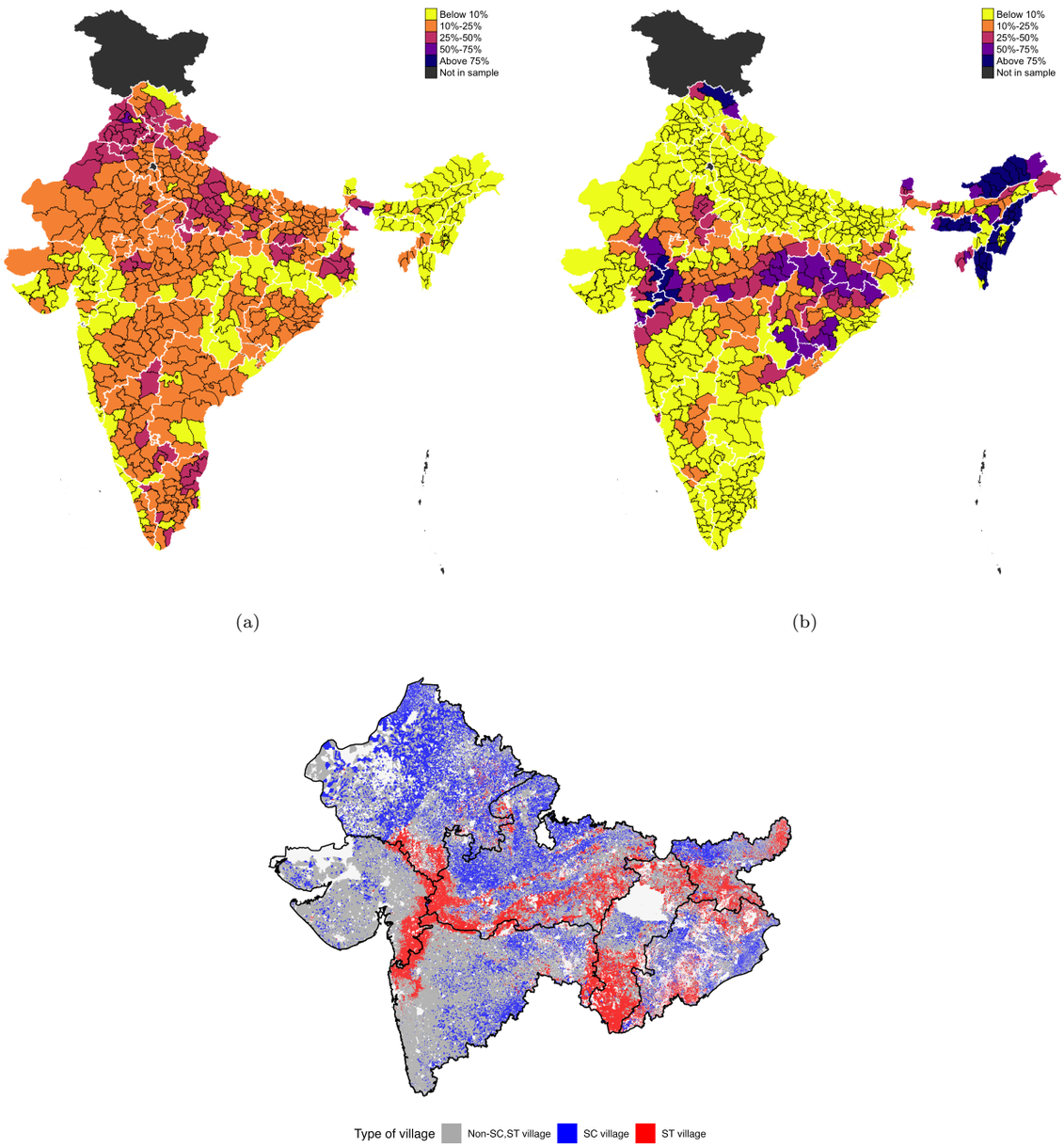
and the distribution of drinking water infrastructure. Unlike existing analyses conducted at district and state levels, our village-level analysis reveals the extreme concentration of Scheduled Tribe communities within villages, which we use to identify and measure the association with access to drinking water structures. Scheduled Tribe populations are among the most marginalized in India, and yet, are rarely the foci of even literature studying the role of marginalization in public resource distribution. Second, we analyse the role of aquifer characteristics in determining the suitability of existing drinking water infrastructure for serving current and future domestic demands. Although aquifer characteristics are among the greatest determinants of pumped water volumes, they are understudied in the context of drinking water supply [5, 12]. On the other hand, government policies and targets rarely consider the variation in the types, scales and costs of drinking water infrastructure required to serve populations residing across different types of aquifers. Finally, we unpack the role of pumping energy in drinking water access. Pumping energy is one of two essential inputs required in the provision of reliable and safe drinking water supply. While rural drinking water literature largely overlooks the role of pumping energy, literature on rural electrification finds recently built electricity infrastructure to be incapable of serving high power needs required for water pumping [13–15]. We differentiate the various drinking water structures based on pumping energy needs and, study the role of rural electrification in shaping their respective distributions.

Scheduled tribe communities are among the most marginalized and least studied in India

The impact of ethnic divisions and concentration of marginalized population on the distribution of public resources has been widely studied across different geographic contexts including parts of Africa, the United States, India, Brazil and Indonesia [7, 8, 10, 16, 17]. In India, it has largely been studied for two major marginalized communities of Scheduled Caste (SC) and Scheduled Tribes (ST), and to a smaller extent for the Muslim community.[2]. However, treatment of SC and ST communities differs widely in this literature. While STs are often acknowledged as the most marginalized, they are largely lumped with SCs, used as a comparison group representative of 'weakest' political representation, or left out entirely. On the other hand, the same studies treat SCs with nuanced and varied lenses of numerical dominance, sub-caste divisions and neighborhood-level segregation [11, 16, 25–30]. This contrasting treatment likely stems from lack of granular data combined with the relative geographic isolation of STs near forested areas making it hard to identify them at broad geographic levels [9, 16, 27, 29]. Outside the north-eastern Indian states, where almost entire populations are tribal, ST households comprise 10% or more of the total population only in seven states spread across the central belt. SC populations on the other hand are more evenly spread throughout the country (figure 1). We restrict our analyses to these seven states with sizable tribal populations and focus on village outcomes to isolate the associations with ST concentration.

Although the geographical concentration of ST households is apparent at the state and district-levels (figure 1(b)), employing granular village-level data, we find that in fact concentration of ST population stems from village-levels. figure 2 illustrates the distribution of ST and SC populations across villages in varying proportions, where x-axis denotes village-level and y-axis state-level population proportions of STs and SCs. Across the seven states, the greatest proportion of STs reside in villages where they constitute 90% or greater of the village populations. Whereas, majority of the SC populations reside in villages where they make up 10%-20% of village populations. The village-level concentration of ST population has important implications for access to public resources, since villages are the smallest unit of nearly all government programmess in India. For instance, drinking water and rural electrification, the two public resources included in our analyses, began with village based targets before expanding to universal household coverage. Even today, despite the household-level targets of electricity access and drinking water supply, the physical infrastructure responsible for their respective deliveries is

[2]Scheduled Caste or dalits occupy the lowest rung of the Hindu Varna (caste) system which stratifies society into heirarchical categories. Whereas, Scheduled Tribes is a term first coined by the British colonists to identify indigenous people living in relative isolation near forested land [18, 19]. There is rich literature challenging distinction between the two groups, and the relative isolation of Scheduled Tribes [20, 21]. Whereas, the strong association of Scheduled Tribes with forested land, and their subsequent dispossession, and the latter's despoliation is widely recognized and deeply studied [22–24]



(a) (b)

(c)

Figure 1: District-level distribution of (a) Scheduled Caste and (b) Scheduled Tribe population across India and (c) village classification as ST, SC, or non-SC/ST Data is based on Census 2011, taken from SHRUG [31]. Black line represent district boundaries and white lines represent state boundaries in (a) and (b); black lines represent state boundaries in (c).

still largely village-wide i.e. transformers for electricity and handpumps or wells for water. We treat villages with concentrations of ST population of at least 80% as ST villages and those with at least 20% SC population as SC villages. Overall, approximately 20% of villages are identified as ST, 26% as SC, with a small overlap of 0.1% of villages which are identified as both ST and SC, we drop the overlapping villages from our sample.

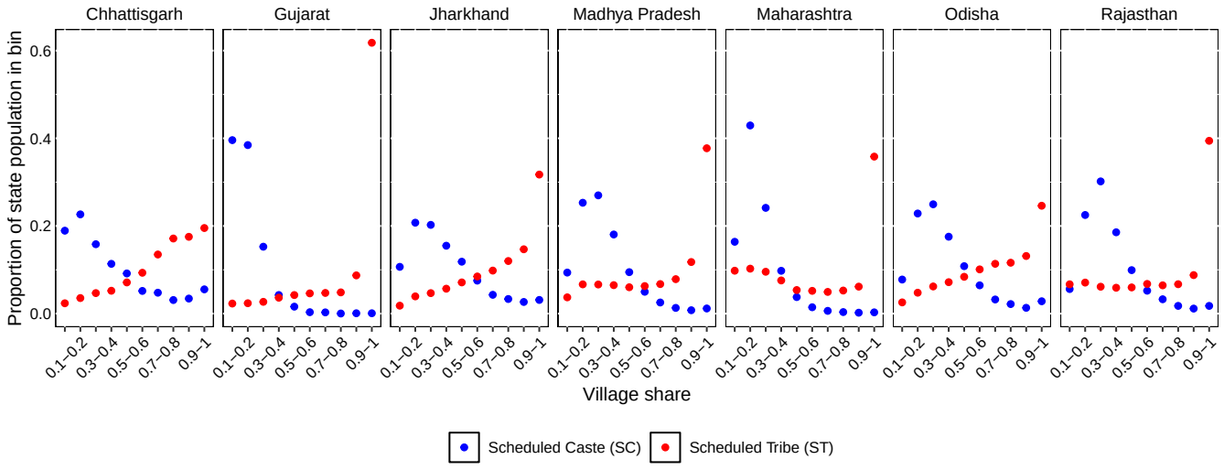


Figure 2: Distribution of Scheduled Caste (SC) and Scheduled Tribe (ST) village-level populations in 2011 for the seven states in our sample. Data taken from SHRUG [31] (n=235,444 villages). X-axis represents deciles of villages population shares of SC and ST. Y-axis represents the proportion of the state population in the respective village share bin.

Moving targets and persisting inequities in drinking water coverage

The Accelerated Rural Water Supply Programme (ARWSP) introduced in the early 1970s provided a set of targets for the roll-out of drinking water infrastructure across rural India. The programme was the first to introduce numerical targets, which included 40l of drinking water per capita per day and access to a minimum of one handpump or a public tap for every 250 people at a distance of no more than 1.6km [32]. Since its introduction, the ARWSP has gone through many iterations with two lasting impacts - first the structure of the numerical targets has remained the same with progressive changes made to the volume, source and distance criteria. For instance, a 2013 version mandates 55l per capita per day requirement via piped water supply systems within 100m of the household (or 30 minutes of fetching time) [33]. Second, handpumps and to a lesser extent, public taps have remained the water structures of choice for the expansion [33]. Even in 2011, handpumps were the most common source of drinking water in our sample, with an average of 43% households reporting using handpumps as their primary drinking water source (figure 3(a)). The remaining water structures include uncovered and covered wells, taps, tubewells, and just over 5% of households cumulatively using surface water based sources such as springs, rivers, lakes etc.

The different groundwater-sourced water structures are largely based on two types of wells - dug wells and bore wells. Bore wells are constructed using drilling machines and can be as deep as required for accessing water to the order of thousands of meters (Karnataka's deep drinking water wells are one such example [34]). Handpumps, tubewells and taps all require bore wells, but greatly vary in water productivity, as handpumps rely on manual pumping compared to mechanized pumping in tubewells and taps. While large pumpsets compensate deep water levels among taps and tubewells, handpumps on average are functional only up to 50m. Dug wells are constructed using manual or animal labour, and therefore, limited to depths of 12m on average [35]. Water structures and whether they are based on mechanized pumping influence rates of extraction and thereby drinking water coverage. We examine the distribution of, and estimate the extraction volumes from handpumps, uncovered wells, taps and tubewells. We assume handpumps and uncovered wells to be based on manual modes of extraction, and taps and tubewells on mechanized pumping. We refer to taps and tubewells cumulatively as pumped water sources.

Besides depths and the presence of pumping energy, productivities of water structures are also influenced by the aquifer characteristics of the inter-laying rocks between the land surface and groundwater sources. There is growing recognition of their role in determining suitability of groundwater based drinking water supply, particularly in hard-rock aquifers which cover nearly two-thirds of the country's landmass [36, 37]. However, aquifer characteristics are rarely considered in research that investigates and informs the distribution of drinking water access in India. Indian drinking water literature largely focuses on the distribution of tap water (despite its relative low presence), and the role of caste and religion in determining the same [38–41]. We find a distinct socioeconomic pattern to the distribution of water structures which motivates our analysis (figure 3(a)). While close to 27% of households on average reported taps and tubewells as their primary source of drinking water, the proportion was less than half of that at approximately 11% among ST villages on average. Instead, over a quarter of households on average relied on uncovered wells for their drinking water supply in ST villages.

Pumping energy: the missing piece in drinking water supply

Pumping energy is a key component of groundwater abstraction, evidenced by the vast pumping energy-groundwater irrigation nexus literature in India [42–48]. Yet, it is scarcely mentioned in rural drinking water supply despite the fact that over 85% of all drinking water sources in rural India are based on groundwater - one reason could be the relatively low levels of pumped water sources in existing drinking water structures [49]. Mechanized pumping is required to operate in only two of the five categories of drinking water structures recorded by the population census of India. In reality, pumps can be, and are often used across all types of drinking water structures. They reduce drudgery, ensure consistent flow and are essential for piped water supply. Pumping energy is essential for village-based drinking water structures only in regions where water depths exceed those accessible by handpumps of approximately 50m (UNICEF). Water at deeper levels require submersible pumps for

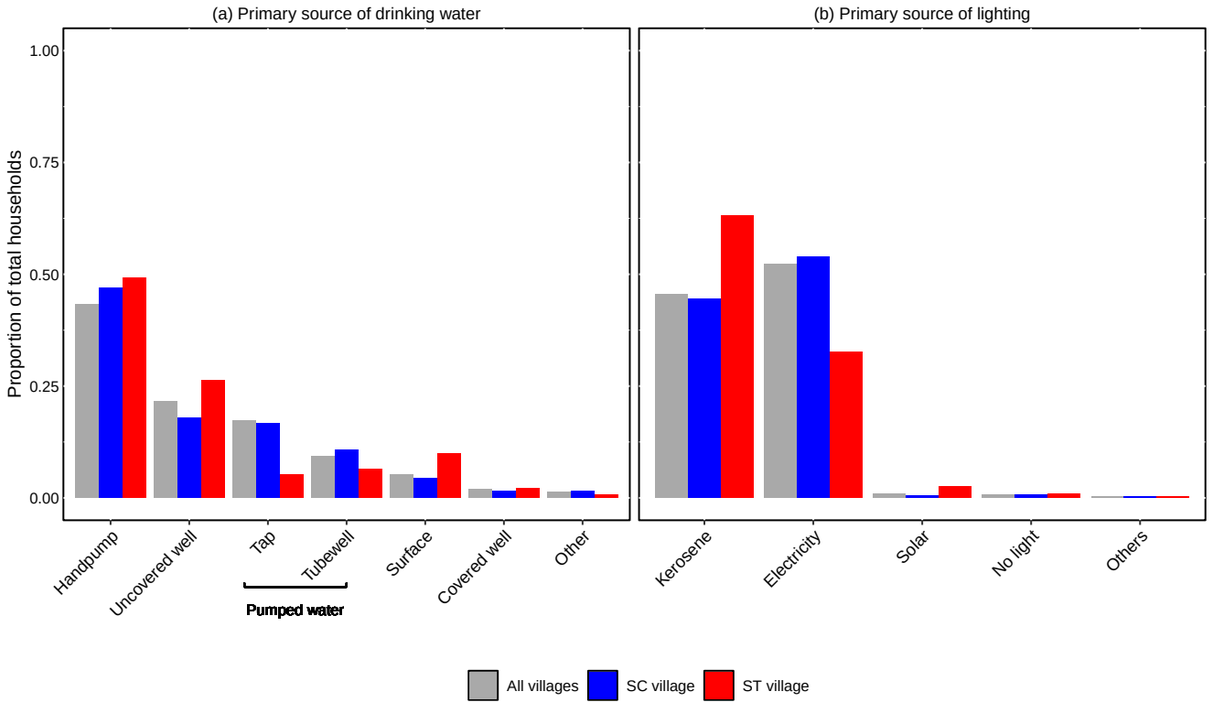


Figure 3: Distribution of sources of drinking water and lighting across SC villages ($\geq$ 20% SC) ST villages ($\geq$ 80% ST), and all villages in our sample. Y-axis represents the average proportion of households in a village reporting utilising the respective source as their primary source of (a) drinking water or (b) lighting. Data compiled using Census 2011 (n=235,444), full details can be found in the Methods section.

extraction which can only be run on electricity.

On the face of it, electricity is an unlikely constraint in drinking water supply due to the tremendous expansion of rural electrification across India in the last decade (see *Saubhagya* dashboard). However, the focus of rural electrification has been largely low power household-based supply since the early 2000s with limited expansion in high capacity uses such as submersible pumps [13, 15]. As a result, adequacy of current water supply in newly electrified regions with deep water levels remains doubtful in the absence of additional electric infrastructure upgrades. Key policy documents on rural drinking water too fail to explicitly mention pumping energy is an essential input [33, 50].

An important manifestation of this oversight is in recorded data - we found no government collected data on energy used in drinking water structures. In its stead, we use household electrification as a proxy for the presence of electricity infrastructure across villages to study the role of electricity in drinking water provision. Household electrification is a weak proxy as it does not ensure the presence of electricity infrastructure required to power submersible pumps, but it is the only reliable data on electricity available at the village-level. On average, a little over half of all households reported using electricity as their primary light source, the remaining largely used kerosene. The average proportion of household using electricity was less than a third among ST villages (figure 3 (b)). We use satellite based nightlight data as a supplementary measure of electrification.

Results

Villages with high concentrations of ST population fare worse for both electricity and pumped water access

We begin by looking at the associations between concentrations of ST and SC within villages, and their respective distribution of drinking water supply and lighting sources (figures 4 and 5). We consider three major sources of drinking water - handpumps, uncovered wells and pumped water (taps and tubewells combined). Together, the three sources accounted for over 91% of village average household water supply in 2011. We find ST villages on average to be associated with 6 percentage points (pp) fewer households using pumped water sources for their drinking needs. The negative penalty remains even after controlling for the proportion of households using electricity, suggesting the presence of constraints beyond electricity infrastructure in using pumped water infrastructure. The lower use of pumped water is largely compensated by uncovered wells. On average, over 2pp more households report sourcing drinking water from uncovered wells in ST villages. SC villages on the other hand are modestly better off, with roughly 1pp and nearly 2pp more households using pumped water and handpumps respectively. We present the full set of results with all sources of drinking water in Appendix A1.

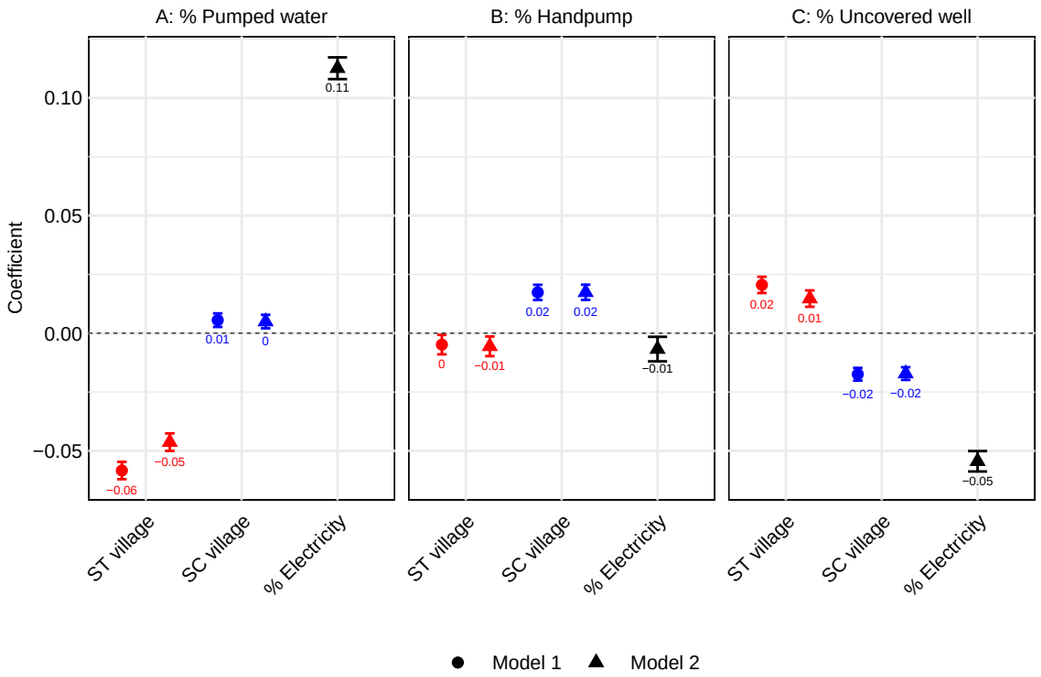


Figure 4: Coefficient plots for linear regressions of the proportion of households using pumped water (Panel A), uncovered wells (Panel B), and handpumps (Panel C) as the primary source of drinking water on the presence of ST village binary estimated using equation 1 and dataset constructed using SHRUG [31] and Census 2011 data. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. We control for the non-ST/SC population, total number of households, presence of roads, distance to district HQ (km), and village area (hectares) in all models. We also present a second set of models (Model 2 coefficients) in which we control for electricity availability - % electricity is the proportion of households that use electricity as the primary source of lighting. Standard errors are clustered at the district level. Significance codes: ***: 0.01, **: 0.05, *: 0.1. Full tabular results are in A1

On average, approximately 27% and 21% of all households at the village-level sourced their drinking water from pumped sources and uncovered wells in 2011 respectively. The effect sizes therefore translate to a 22% decrease in the usage of pumped sources, and a 10% increase in uncovered wells among households in ST villages on average. Our results for ST population are robust to alternate model specifications where we use the proportion of ST and SC population instead of the binary indicators (appendix Table A6). However the small positive association with SC population is lost for households using pumped water.

Looking at primary sources of lighting, ST villages are penalized to a greater extent. We find ST villages to house 11pp fewer households using electricity, whereas SC villages are modestly better off by approximately 1pp more households using electricity (Table A2). Unsurprisingly, results for kerosene operate in the opposite direction since it is the most popular substitute for fueling lights. On average, 52% households in our sample reported using electricity and 45% households using kerosene in 2011 at the village-level implying an average loss of 21% households using electricity among ST villages. Our results are robust to alternate model specifications only in the case of ST villages, where we use the proportion of ST and SC population instead of the binary indicators (appendix Table A7). Interestingly, proportion of households using electricity is negatively associated with the proportion of SC population, suggesting the large effect of villages without any SC population in the overall association (appendix figure A1). We also consider night time luminosity captured by the Defense Meteorological Satellite Program (DMSP) Operation Linescan System (OLS) as a way to triangulate data on electricity use. Column (3) of Table A2 reports the percent change in night time luminosity of villages averaged over 2010-2012. The results are similar qualitatively, where ST villages are associated with approximately 11% decrease and SC villages with approximately 3% increase in night-time luminosities.

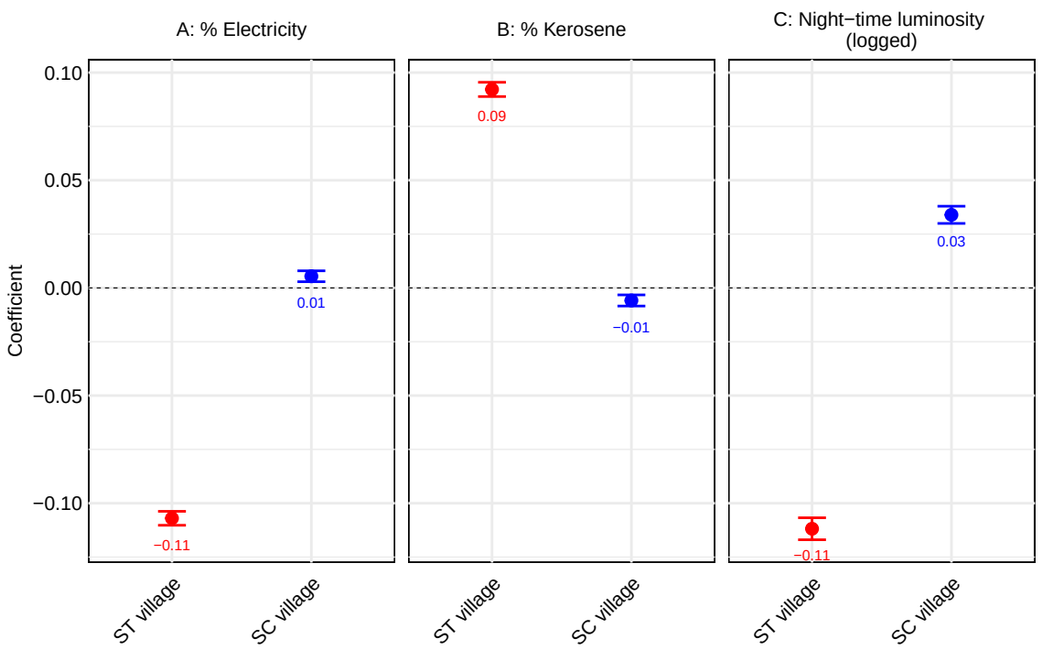
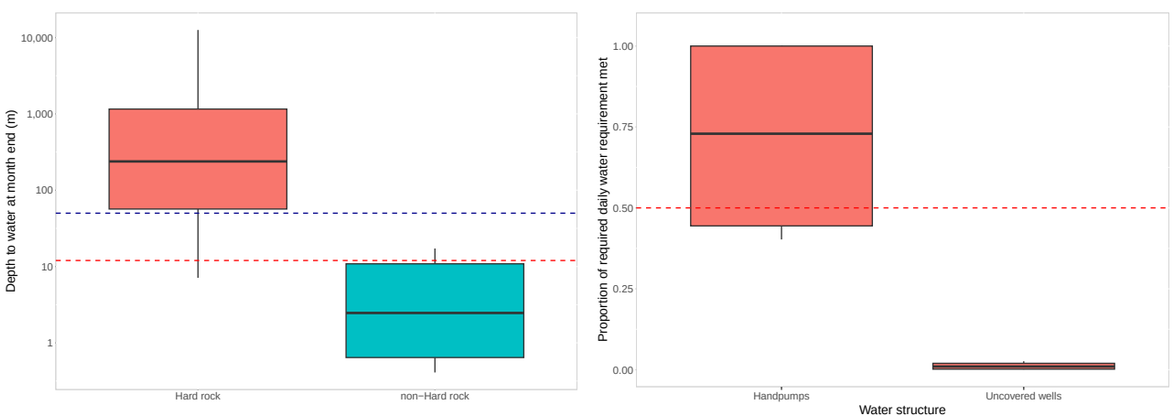


Figure 5: Coefficient plots for linear regressions of the proportion of households using electricity (Panel A) and kerosene (Panel B) as primary sources of lighting, both taken from Census 2011 data. The outcome in Panel C is the night time luminosity of villages, averaged over 2010 to 2012. All models are linear regressions of the outcome on the presence of ST and SC village binaries estimated using equation 1 and dataset constructed from SHRUG and Census data [31]. ST village is a binary indicator for $\geq 80\%$ ST population and SC village indicates $\geq 20\%$ SC population. We control for the non-ST/SC population, total number of households, presence of roads, distance to district HQ, and village area. Standard errors are clustered at the district level. Significance codes: ***: 0.01, **: 0.05, *: 0.1. Full tabular results are in A2

Hydrogeologies play an important role in determining drinking water availability

Productivity of groundwater structures can vary widely depending upon aquifer characteristics, permeability, groundwater stock, presence of faults or fractures and, depth to water in addition to water infrastructure related characteristics [51–53]. Despite the near ubiquitous presence of groundwater-based drinking water infrastructure across Indian rural households, the actual availability of water is likely to vary dramatically across space and time. Aquifer characteristics governing drawdowns are among the most important in determining pumped volumes, and therefore, also in predicting long-term sustainability of drinking water infrastructure [12, 52, 53]. Drawdowns are temporary drops in water-levels caused by continuous pumping and can lead to temporary water failures when they exceed depths of respective water structures. For instance, if drawdowns exceed 50m due to a dry spell or overpumping, a handpump capable of lifting water from a maximum of 50m will be rendered non-functional until the water level recovers.

We examine the respective suitabilities of the different water structures across one hydrogeology each from hard rock (granite) and non-hard rock (alluvial) categories. Since aquifer characteristics are location specific, we choose aquifer characteristics of Transmissivity and Storativity in one district in Odisha to simulate drawdowns resulting from meeting daily domestic water needs for a village with 750 residents. Simulated drawdowns allow us to evaluate the suitability of two of the three commonly used water structures - handpumps and uncovered wells - in extracting the requisite daily volumes of water. Tubewells and taps can lift groundwater from greater depths as they rely on mechanized pumping and are thus not similarly constrained by drawdowns subject to pump-set capacity and electricity supply.



(a) (b)

Figure 6: (a): Drawdowns at the end of a month across Granite and Alluvium hydrogeologies in Mayurbhanj based on 55 litres per capita requirement for an average of 750 people village. Simulated depths are based on Equation 2 in the Methods section. The red dotted line denotes an average depth of uncovered wells of 12 m and blued dotted line denotes average depth of deep handpumps of 50m. (b): Proportions of required daily drinking water met by handpumps (max depth of 50m) and uncovered wells (max depth of 12m) in hard rock villages. Daily volumes are fully met in non-hard villages by both uncovered wells and handpumps. Simulated volumes are based on Equation 2 in the Methods section. The red dotted line denotes 50% of the total daily requirement.

Drawdowns in non-hard rock hydrogeologies rarely go beyond tens of meters compared to the exponentially deeper levels water plummets to in hard rock regions - the median drawdown for hard rock villages was nearly 300m compared to less than 5m for non-hard rock (figure 6(a)). Consequently in non-hard rock villages, non-mechanized uncovered wells are sufficient in meeting drinking water needs, albeit labor intensive. Whereas, deep drawdowns limit the volumes of water that can be pumped from non-mechanized sources in hard rock villages (figure 6(b)).[3] Even handpumps built for deep aquifers and operational up to 50m are likely to meet only 50% of the daily water needs, while uncovered wells with an average depth of 12m are likely to provide less than 1% of the assumed daily requirement (figure 6(b)). In contrast, a tubewell of approximately 100m depth and powered by 1 HP pumpset is likely to fulfill daily domestic water requirements in hard rock regions.

[3]Our simulations are simplified estimations assuming uniform T and S values for the entire range of depths of the water infrastructure. In reality, these values vary across depth influencing drawdowns and water productivity. We also do not consider recharge through precipitation, which is likely to be higher in uncovered wells on account of direct capture.

Hydrogeologies do not explain the distributional differences between high and low concentrations of marginalized population villages

While hydrogeologies play an important role in determining success of water infrastructure, it is unclear whether they also play a role in determining the types of water structures that get built across villages. We find hard rock areas to have greater pumping energy needs owing to the deep drawdowns caused by continuous pumping. Hence, we expect greater proportions of households in hard rock dominated areas to source drinking water from pumped sources than in non-hard rock areas. In Mayurbhanj in Odisha, however, a smaller proportion of households in granite-dominated villages sourced drinking water from pumped water than in alluvial-dominated villages in 2011 (appendix table A10). As both socioeconomic and biophysical factors may shape access to drinking water resources, we examine whether marginalized populations are concentrated within distinct hydrogeologies, and parse out the associations between hydrogeologies and concentration of marginalized populations and access to drinking water and lighting. We classify villages as hard or non-hard rock villages using district geological maps in Odisha, and find ST concentration to to be greater in hard-rock areas (figure 7).

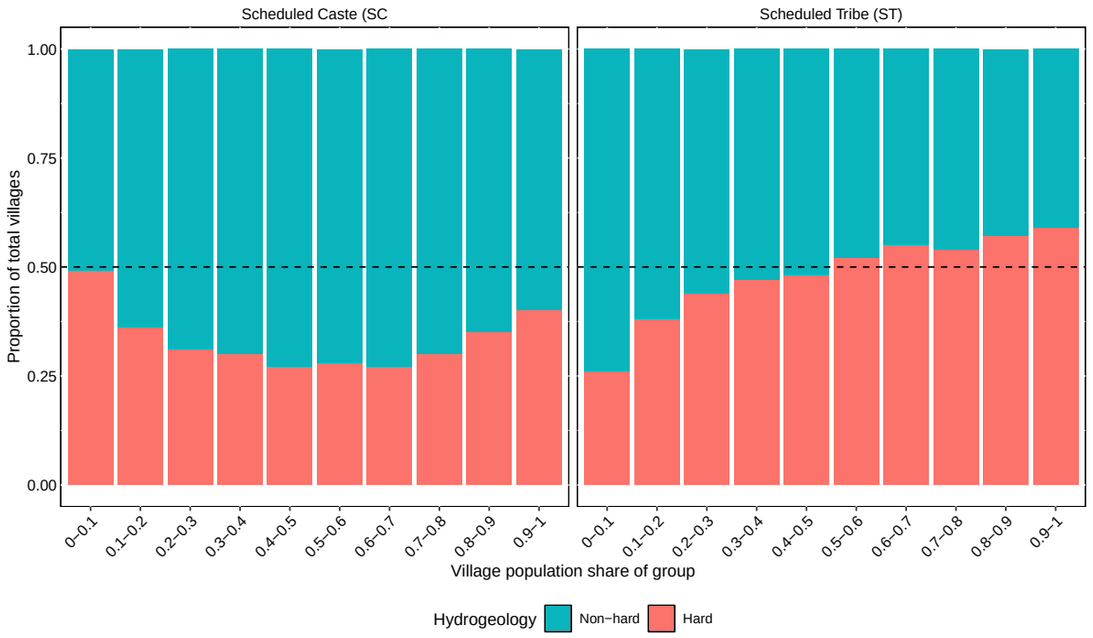


Figure 7: Village-level distribution of hard and non-hard hydrogeologies across deciles of Scheduled Caste (SC) and Scheduled Tribe (ST) share in villages. Each bar represents the proportion of hard and non-hard villages in the respective village population share decile. Villages are classified as hard-rocks if a majority of the land area covered by rocks constitutes hard rock. Data compiled by authors (n = 44,661), full details can be found in the Methods section.

If hydrogeologies were duly considered while constructing water structures, we would expect ST villages to have a greater proportion of households utilising pumped water sources than non-ST villages. However, we find no evidence of this - ST villages are associated with over 3pp fewer households using pumped water, and 11pp fewer households using electricity compared to non-ST villages (figure 8). Electricity is a key source of pumping energy and may be the limiting factor for access to pumped water. However, even after controlling for the proportion of households using electricity, ST villages face a penalty of close to 3pp in accessing pumped water. On average, only 25% households reported pumped water access among villages in Odisha translating to a loss of 12% households with pumped water supply among ST villages. Hydrogeology does not appear to worsen or alleviate their penalty, implying the distribution of drinking water structures does not consider specific hydrogeologies of locations despite its disproportionate influence on water productivity. We observed no similar penalty for SC villages.

Interestingly, 3pp fewer households use electricity in hard rock villages than in non-hard rock villages on average, implying the concentration of marginalized population in hard rock areas may be driving the difference (figure 8b). The lack of electricity among ST villages in hard rock areas also suggests the urgent need for, and additional hurdles to providing pumped water infrastructure to these villages.

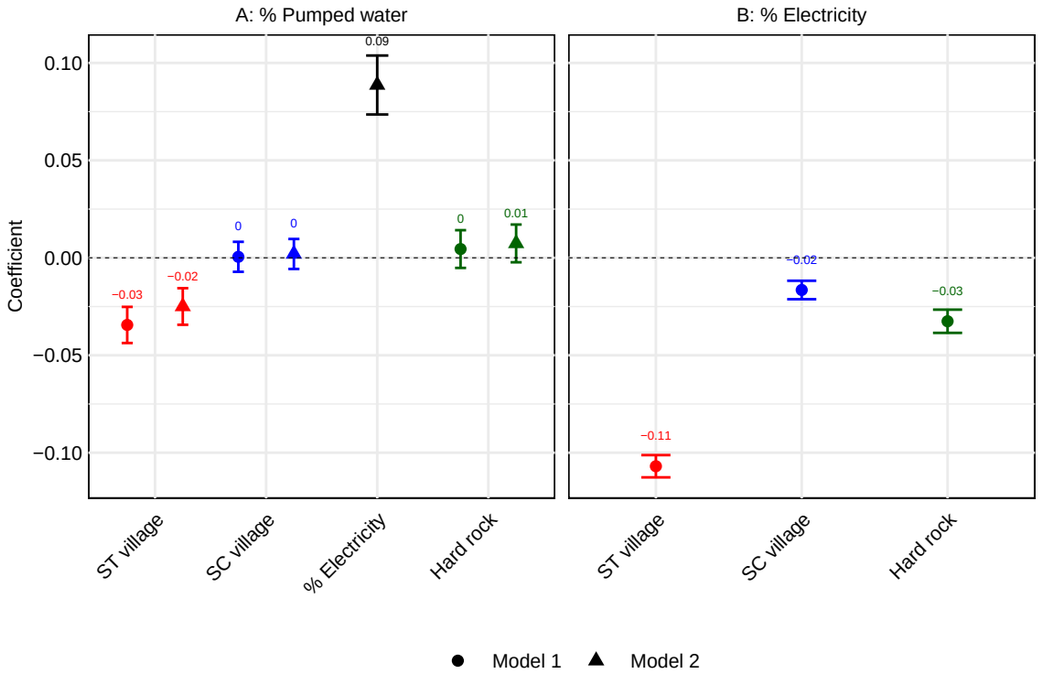
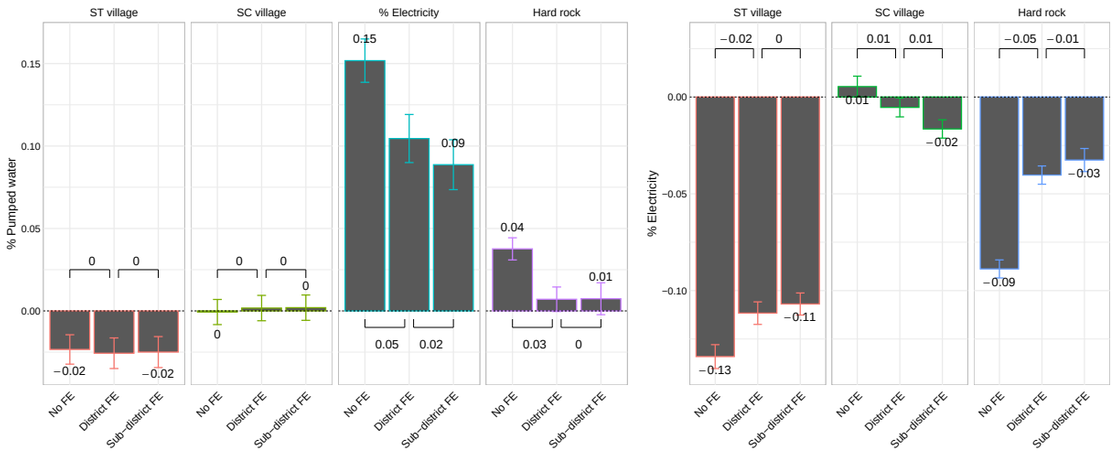


Figure 8: Coefficient plots for linear regressions of the proportion of households using pumped water (Panel A) as the primary source of drinking water and electricity (Panel B) as the primary source of lighting on ST, SC village binaries and type of rock estimated using 8 and data from Census 2011, SHRUG[31], and geological maps. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. We control for the non-ST/SC population, total number of households, presence of roads, distance to district HQ (km), and village area (hectares) in all models. In Panel A, we also control for electricity availability (Model 2 coefficients) - % electricity is the proportion of households that use electricity as the primary source of lighting. Standard errors are clustered at the district level. Significance codes: ***: 0.01, **: 0.05, *: 0.1. Full tabular results are in appendix A3

Although hydrogeologies appear to have a weak association with the distribution of village-level water infrastructure, it may be that they are considered at district-level or state-level planning (as is the case in several government programmes targeting lagging districts and states [54, 55]). A village-level comparison within sub-districts is likely unable to capture such targeting. To examine the relative distributional inequities across different administrative levels, we follow a method proposed by Asher et al [30] (see Methods).

Overall, we find hydrogeologies to have little association with the distribution of both pumped water and electricity and largely do not explain the ST village penalty (figure 9). Districts are the only level where we find hard-rock villages on average to be associated with greater proportions of households sourcing drinking water from pumped sources (figure 9(a)). However, no similar statistically significant difference is present among sub-districts or villages which have greater or lower presence of hard-rock hydrogeologies. Whereas, we find ST village penalty at every administrative level, with the villages being worst off for both pumped water and electricity. In fact the ST village penalty in case of electricity is quite stark where almost the entire effect is captured at the village-level (figure 9(b)). On average, there were 13pp fewer households using electricity in ST villages in Odisha after controlling for the dominant hydrogeology. This total effect is largely captured by village-level effect of close to 11pp, which translates to an average loss of 36% electrified households among ST villages. For SC villages the distributional inequity across all geographic levels is largely missing for pumped water, and are small and mixed for access to electricity.



(a) (b)

Figure 9: Coefficient plots of regressions results of proportion of population with (a) and sources of drinking water and (a) electricity on the presence of hard rock, ST, SC village binaries and hard rock in Odisha based on Equation 8 Full set of results are provided in appendix A4 and A5. The first bar is from the model without fixed effects and represents the 'Total effect' i.e. effect across villages estimated using a model without fixed effects. This is decomposed into the District, Sub-district, and Village effects. District effect ($\alpha_{District}$is the effect across districts, calculated by subtracting the effect across villages within districts (using $\beta_{District}$ from the district fixed effects model) from the effect across villages (β_{Total} i.e. total effect). Similarly, the Sub-district effect ($\alpha_{Sub-district}$) is calculated as $\beta_{District}$ (coefficient from district fixed effects model) - $\beta_{Sub-district}$ (coefficient from sub-district fixed effects model). Lastly, the Village effect ($\alpha_{Village} = \beta_{Sub-district}$) captures variation across villages within sub-district and is the coefficient from the sub-district fixed-effects models. We control for the non-ST/SC population, total number of households, presence of roads, distance to district HQ (km), and village area (hectares) in all models. We alo control for electricity availability - % electricity is the proportion of households that use electricity as the primary source of lighting - in (a). Standard errors are clustered at the district level. Significance codes: ***: 0.01, **: 0.05, *: 0.1. Full tabular results are in appendix A4 and appendix A5

Discussion

The Indian government has been increasing rural drinking water coverage through various national-level programmes with the ultimate aim of providing universal household coverage of piped-water supply. However, the most recent statistic available from 2018 reveals a large gap in ambition and reality, with a mere 17% of rural households with piped-water connections [56]. Indian rural households access drinking water majorly from groundwater based sources through wells, handpumps and taps. These water structures greatly differ in terms of physically and economically accessible volumes, and water quality. Large variations exist across different geographies and seasons for a single water structure even in the presence of sufficient water volumes. The role of biophysical characteristics in determining water productivity is widely accepted, yet their role in perpetuating regional inequities and interaction with socioeconomic data remains largely unexamined. We study the distributive inequities of existing water infrastructure for marginalized populations, and estimate the likelihoods of the water structures in meeting daily drinking water demands by employing a novel village-level dataset including biophysical and socioeconomic data.

We consider the two largest groups of marginalized population in India - Scheduled Castes (SC) and Scheduled Tribes (ST), and focus our analysis on seven Indian states where STs comprise over 10% of the respective state populations. By doing so, we are able to identify their extreme geographic concentration within villages, which is in direct contrast to the distribution of the SC population. Nearly 50% of the ST population across the seven states resided in villages where they made up over 80% of the village population, while less than 3% of SC population resided in similarly concentrated villages. Our study addresses an important oversight in existing literature on the distribution of public resources among marginalized population by revealing this extreme concentration of ST populations. We find ST-concentrated villages to experience an average decrease of 22% households sourcing drinking water from pumped groundwater sources. This decrease is largely compensated by an increase of 10% in open unprotected wells. Lower levels of pumped water in ST villages could be a result of missing electricity infrastructure. Indeed, we find ST villages are associated with 11pp fewer electrified households as compared non-ST villages. Drinking water and electricity are two of many such public resources which are distributed through village-level infrastructure. These village-level infrastructure - water structures such as wells and handpumps in case of drinking water and transformers for electricity - are critical determinants of supply adequacy. While our analysis is not causal, we identify ST concentrated villages as a possible lever to enhance poverty targeting of government programmes and help plug subsidy leakages.

Aquifer properties are widely recognized as important factors determining extractable water volumes, and yet, are largely missing from rural drinking water literature [12, 52]. We find non-motorized water structures to be largely inadequate in furnishing domestic water needs across hard-rock villages. Deep drawdowns in these regions makes electrified pumping of water essential. Even handpumps which operate up to a maximum depth of 50m are only capable of furnishing less than half the daily volume for a village with an average population of 750. However, we are unable to consider multiple water structures which could be present in one village, since the data is not collected. While multiple handpumps could meet daily water needs of an entire village, the same is unlikely to be true for uncovered wells. In addition to being unprotected from contamination, uncovered wells with an average depth of 12 m are likely capable of serving only a maximum of 20 people each day, implying the need for numerous such wells for a village with 750 residents.

Despite their role in determining water productivity, we find aquifer properties to play a small role in distributional pattern of drinking water structures, especially for ST villages. In our dataset, greater proportions of ST populations reside in hard-rock villages. Yet, the negative association between ST villages and pumped water does not diminish even after considering the presence of hard-rocks. ST villages are thus doubly penalized - not only are they more likely to be in hard rock regions where mechanized pumping is essential due to deep drawdowns, they are also more likely to have less productive water structures such as uncovered wells. On the other hand, wells of approximately 100m depth and electricity capable of supporting a 1HP pump are adequate water infrastructure components to serve domestic water needs of hard-rock villages on average.

Overall our results paint a grim picture of the plight of ST villages similar to existing literature. However, our

results imply that geographic segregation of STs has biophysical dimensions in addition to socioeconomic ones. Both these dimensions are important determinants of drinking water access. In contrast to literature suggesting ST penalties stemming from low political representation, we find equally poor outcomes in Chhattisgarh and Jharkhand, states expressly created to serve ST interests [26]. Social homogeneity is widely posited as an aid in bottom-up bargaining for social goods, but fails to explain the poor outcomes of ST-concentrated villages [26, 38].

While we do not offer reasons behind the poor performance of ST communities, our results have important implications for remedying their situation. We find the greatest distributive inequity to exist at the village-level making it an important site for targeting affirmative government policies. Additionally, it is imperative to consider local hydrogeological characteristics explicitly in drinking water policies and planning. Our results imply that water planning based only on population parameters is unlikely to succeed in hard rock villages with deeper drawdowns, greater pumping energy requirements, and with higher probabilities of seasonal failures. Since ST populations are likely to be concentrated in these regions, they are also most likely to be impacted. Besides identifying the urgent need for pumped water sources among hard rock areas, we are unable to provide a comprehensive analysis of water structure requirements in view of seasonal and annual fluctuations in groundwater level, precipitation and added uncertainty on account of climate change. Electricity in service of pumping energy is another important factor that needs to be brought front and center in water planning. While ST villages are doubly penalized with lower access to pumped water and electricity, their concentrations in hard rock regions and within villages, offers a targeted route to improve their situation. Future policies and research need to consider biophysical elements impacting public resource access, and identify areas of overlap with the distribution of marginalized populations. These overlaps while creating unforeseen poverty traps for marginalized communities, could also offer opportunities of compounding positive impacts of public resources.

Data

We use three main sources of data for our analyses - Population census data, night-time luminosity and forest cover compiled by Asher et al [31]; Primary sources of drinking water and lighting from the house-listing tables of the population census data in 2011 from data.gov.in and; hydrogeological data for Odisha from the state's Centre of Environmental Information System (ENVIS). We match all the datasets at the village-level.

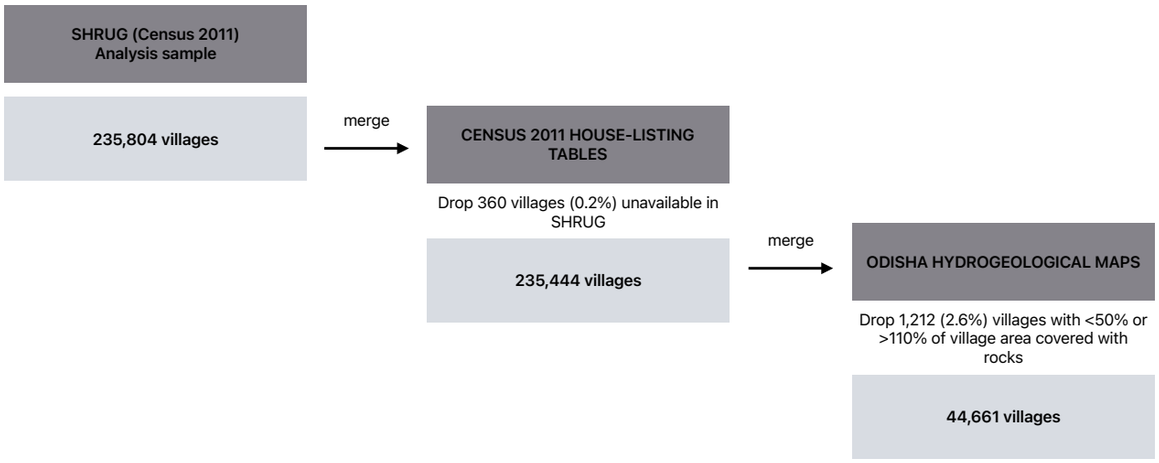


Figure 10: Combining data from SHRUG, Census 2011 house-listing tables, and Odisha hydrogeological maps

SHRUG

We use three types of data from v2.0 of the Socioeconomic High-resolution Rural-Urban Geographic Platform for India (SHRUG). SHRUG provides data at the "shrid" level, which is a geographic unit with stable boundaries which occasionally combines multiple villages. We limit our analysis to shrids that constitute a single village and hence, present all results at the village level.

Village Amenities modules - We use the Village Amenities modules from the 1991, 2001, and 2011 Censuses of India to calculate the village population share by marginalized-group and to control for various village characteristics[4] The census reports the total population, number of Scheduled Tribes (STs), and number of Scheduled Castes (SCs) for each village. Using this, we calculate the ST and SC share for each village. We define ST villages as those that have $\geq 80\%$ ST share in at least one of the three years and $> 50\%$ ST share in all three years. ST village is thus a time invariant binary variable that identifies villages with at least a majority share of STs over three decades. Similarly, we define SC villages as those that have $\geq 20\%$ SC share in at least one of the three years and $> 10\%$ SC share in all three years.

We also define village-level controls using this data - village area (in hectares), non-ST/SC population (total - ST - SC), and total number of households. We control for remoteness of villages in two ways - road (presence of pucca or black-topped roads) and distance to the district headquarters (in km)

Night time luminosity - We use the night time luminosity module provided in SHRUG [31]. We use annual mean night time luminosity as an additional proxy for village electrification. Night time luminosity is captured by the Defense Meteorological Satellite Program Operation Linescan System (DMSP-OLS), expressed as the average brightness of all pixels in the village boundary, and ranges from 0 to 63. We use village-level night time luminosity averaged across 2010, 2011, and 2012.

Forest cover - We use the forest cover module to identify forested area [31, 57]. Annual forest cover is measured as Vegetation Continuous Fields (VCF) using satellite imagery which predicts primary and secondary forest coverage for 250m pixels. The satellite images are taken from MODIS, a product which measures tree cover, after which crops and plantations are differentiated from forests using a machine learning algorithm based on

[4]For our sample, the population numbers reported in the Primary Census Abstract are identical to those in the Village Amenities module.

higher resolution satellite imagery. We use percentage area of the village covered by forests averaged over 2001-2020.

Census house-listing data

We combine the census village amenities module with the census house-listing module, which provide proportions of total households accessing the different sources of drinking water (including uncovered wells, handpumps, tube wells, and piped water) for each village [58]. We do this as although the village amenities data provides binary indicators of the presence of various drinking water infrastructures, it does not report how many households rely on each source. Using the house-listing data, we define three outcomes for sources of drinking water - uncovered wells, handpumps, and pumped water (tube wells and piped water, both of which require mechanised pumping). The house-listing data also provide similar breakdowns for the main source of lighting, across kerosene and electricity. We use village crosswalks provided by SHRUG and calculate population weighted averages for units spread across more than one village ($< 0.2\%$ of sample) to combine the house-listing tables with the other census modules. In all we are able to match 99.8% of the villages across census data provided by [31] with the villages in the house-listing module.

Hydrogeological characteristics

We identify primary rock formations in each village by digitising and overlaying geological maps from ENVIS onto village base maps from SHRUG [31, 59] for one of our seven states - Odisha. Digitising and overlaying maps is not precise and combined with the fact that geological boundaries do not follow administrative boundaries of villages results in some villages where we under or over-classify the entire village areas. In such cases, we drop villages that appear to be $< 50\%$ of $> 110\%$ covered by rocks (1,212 villages of the original 45,873). For the remaining 97% villages, we classify villages as hard-rock or non-hard-rock villages - if hard rocks constitute $> 50\%$ of the village area covered by rocks, we classify it as a hard rock village. We define hard rocks based on the range of the transmissivity and storativity values of the rock-types recorded in the district handbooks included in appendix table A11.

Methods

SC/ST villages and access to drinking water and lighting

We use a one-way fixed effects model (Equation 1) to estimate the association between ST, SC concentrations, and access to sources of drinking water and lighting.

$$Y_{ij} = \beta_1 \mathrm{ST}_i + \beta_2 \mathrm{SC}_i + \mathbf{X}_i + \gamma_j + \epsilon_i \tag{1}$$

Where, Y_{ij} is the proportion of households accessing electricity (or kerosene) and pumped sources of water (or handpumps/ uncovered wells) in village i in subdistrict j in 2011. ST_i and SC_i are binary indicators of whether the village is an ST village or an SC village respectively. Our data includes 47,675 ST and 60,227 SC villages. A small proportion of our sample were classified as both ST and SC villages, which we drop from the analysis. Therefore, the comparison group are villages which have ST and SC populations below 80% and 20% respectively.

$\mathbf{X}_i$ is a vector of village controls including the non-SC/ST population, total number of households in village i, binary indicator capturing the presence of road (whether the village has black-topped or pucca roads), distance from district headquarters (km), and village area in hectares. γ_j are subdistrict fixed effects - our estimates therefore represent the differences between ST, SC and non-ST/SC villages *within* sub-districts.

Hydrogeology and access to drinking water and lighting

We conduct two sets of analyses to unpack how varying hydrogeological conditions shape access to water. First, we estimate water volumes that can be extracted from the various drinking water structures across hard and non-hard rock geologies. Second, we use one-way fixed effects regression models to we control for the presence of hard-rock geologies and test whether hydrogeological conditions influence the types of water structures accessed by varying proportions of households across villages.

Role of hydrogeologies in determining extracted volumes from different water structures

We estimate maximum productivities of different drinking water infrastructures based on their respective depths and hydrogeological characteristics of underlying rocks across villages. We restrict our analysis for these simulated results to Odisha's Mayurbhanj district as the greatest proportion of Odisha's ST population reside in Mayurbhanj (16% of the state's population in 2011) and the district is composed of both non-hard and hard rock hydrogeologies. We assume a daily water requirement of 41,250 litres for a representative village of 750 people using the national drinking water policy's goal of 55 litres per capita per day. The 50^{th} and 75^{th} percentile values of village-level populations across Mayurbhanj were 495 and 810 respectively in 2011. We simulate the daily drawdowns of this daily water requirement for bore wells and uncovered wells using equations (2-4).

Drawdowns are the depths to which groundwater levels fall with continuous pumping and subsequently recover once pumping has ceased. We select alluvial and granite as the representative non-hard and hard rock geologies respectively and use their district-average values of Transmissivity and Storativity - geological properties governing drawdowns and recoveries. Average daily water requirements cannot be met without deepening water structures for aquifers where drawdowns exceed depths. We assume well depths of 12m and 50m for uncovered wells and handpumps respectively. For those structures, we run a constrained optimization to calculate new daily pumping rates Q^* which would limit the monthly drawdowns to well depths. Q^*/Q is therefore the proportion of daily water requirement that is met.

For uncovered wells, we assume well radius of 5m and for 50mm for handpumps which are based on bore wells. We further assume that water is extracted for 8 hours a day with a recovery duration of 16 hours during which no water is extracted and pump efficiency of 70%. The drawdowns at the end of a month are a function of the levels to which water drops at the end of the each day and the subsequent increase in water levels during hours of recovery.

Drawdowns at the end of the month -

$$H_s = \frac{Q_1}{4\pi T} W(u_1) + \frac{Q_2 - Q_1}{4\pi T} W(u_2) + ... \quad (2)$$

$$W(u_n) = -0.5772 - ln(u_n) + u_n - \frac{u_n^2}{2.2} + \frac{u_n^3}{3.3} + ... \quad (3)$$

$$u_n = \frac{r^2 S}{4\pi T t_n} \quad (4)$$

Where, H_s is the drawdown at the end of the month, T Transmissivity (m^2/day) and S Storativity. Transmissivity and Storativity are aquifer characteristics. r radius of borehole assumed to be 50mm and t_n time elapsed since pumping rate Q_n start (days).

We also model the energy requirements for pumping water to meet the daily water requirement using the following equations -

1. Daily average energy requirement in month i

$$E_i = \frac{Q_i \rho g H}{3.6 \times 10^6 \eta_{pump}} \quad (5)$$

Where, H is the average depth to groundwater (m), assumed to be 5m at the beginning, to which drawdown from pumping is added. η_{pump} pump-set efficiency and assumed to be 70%. ρ is the density of water (1000 kg/m^3) and g is the gravity constant (9.8 m/s^2).

2. Required Pump Capacity -

$$P_{cap} = \frac{max(E_i) \times 1.341}{t_{max}} \tag{6}$$

Where, $P_{cap,i}$ is the required pump capacity (HP), $maxE_i$ is the maximum energy requirement. t_{max} is the maximum hours of pumping in a day and assumed to be 8 hours.

3. Required hours of pumping -

$$t_i = \frac{E_i \times 1.341}{P_{cap}} \tag{7}$$

Where, t_i is the average daily hours pf pumping in month i.

Role of hard rock geologies in drinking water and lighting access

We use a second set of one-way fixed effects models (Equation 8) to study the role of hydrogeological characteristics in addition to ST, SC concentration in determining access to sources of drinking water.

$$Y_{ijk} = \beta_{Total}\text{ST}_i + \beta_2\text{SC}_i + \alpha\text{Hard}_i + \mu\text{Electricity}_i + \mathbf{X}_i + \epsilon_i \tag{8}$$

As before, Y_{ijk} is the proportion of households accessing electricity (or kerosene) and pumped sources of water (or handpumps/ uncovered wells) in village i in subdistrict j and district k in Odisha during 2011. ST_i and SC_i are binary indicators of SC and ST villages. In addition to village controls of non-SC/ST population, total households, presence of roads, distance from district headquarters, and village area, we add a binary indicator $Hard_i$ to indicate whether the village is situated in a hard-rock hydrogeology or not. To account for the role of electricity in access pumped water sources, we also include proportion of households accessing electricity $Electricity_i$ as a control in models with sources of drinking water as the outcome.

To parse the inequalities in access at different administrative levels, we use equations (9-12) adapted from [30].

$$\alpha_{District} = \beta_{Total} - \beta_{District} \tag{9}$$

$$\alpha_{Sub-district} = \beta_{District} - \beta_{Sub-district} \tag{10}$$

$$\alpha_{Village} = \beta_{Sub-district} \tag{11}$$

$$\alpha_{District} + \alpha_{Sub-district} + \alpha_{Village} = \beta_{Total} \tag{12}$$

We first estimate a model with no fixed effects giving us β_{Total} as the total effect of ST villages on our outcomes of interest. This total effect is an additive composite of the effect of ST villages across villages, sub-districts and districts. Next, we use a district fixed effects models with $\gamma_{District}$, which allows us to estimate $\beta_{District}$, the effect of ST villages within each district. The effect of ST villages across districts can then be calculated by removing the within-district effect, $\beta_{District}$, from the total effect, β_{Total}. Hence, we calculate $\alpha_{District} = \beta_{Total}$ - $\beta_{District}$ as the across-district effect. Similarly, we estimate $\beta_{Sub-district}$ (effect within sub-districts) using sub-district fixed effects $\gamma_{Sub-district}$ to calculate $\alpha_{Sub-district} = \beta_{District}$ - $\beta_{Sub-district}$, which is the effect across sub-districts within districts, and $\alpha_{Village} = \beta_{Sub-district}$ (effect across villages within sub-districts) using sub-district fixed effects. The total effect β_{Total} is then the sum $\alpha_{District} + \alpha_{Sub-district} + \alpha_{Village}$.

References

[1] WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene. Data dashboard, 2022.

[2] Werner Aeschbach-Hertig and Tom Gleeson. Regional strategies for the accelerating global problem of groundwater depletion. *Nature Geoscience*, 5(12):853–861, December 2012.

[3] Tom Gleeson, Yoshihide Wada, Marc F. P. Bierkens, and Ludovicus P. H. van Beek. Water balance of global aquifers revealed by groundwater footprint. *Nature*, 488(7410):197–200, August 2012.

[4] K. J. Joy and Veena Srinivasan. Climate Change and Rural Water Security. In Ravi Agarwal and Omita Goyal, editors, *The Crisis of Climate Change: Weather Report*, pages 149–163. Routledge India, London, 1 edition, September 2021.

[5] Katrina Jessoe. Improved source, improved quality? Demand for drinking water quality in rural India. *Journal of Environmental Economics and Management*, 66(3), 2013.

[6] Sheetal Sekhri. Wells, Water, and Welfare: The Impact of Access to Groundwater on Rural Poverty and Conflict. *American Economic Journal: Applied Economics*, 6(3):76–102, July 2014.

[7] William Easterly and Ross Levine. Africa's Growth Tragedy: Policies and Ethnic Divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250, November 1997.

[8] Alberto Alesina, Reza Baqir, and William Easterly. PUBLIC GOODS AND ETHNIC DIVISIONS. *The Quarterly Journal of Economics*, 114(4):1243–1284, November 1999.

[9] Caste, Community and Collective Action: The Political Economy of Public Good Provision in India.

[10] Yuhki Tajima, Krislert Samphantharak, and Kai Ostwald. Ethnic Segregation and Public Goods: Evidence from Indonesia. *American Political Science Review*, 112(3):637–653, August 2018.

[11] Naveen Bharathi, Deepak Malghan, and Andaleeb Rahman. Ethnic diversity and economic development with spatial segregation. *Economics Letters*, 222, 2023.

[12] Sudatta Ray and Sally M. Benson. Energy models in service of aquifer specific groundwater irrigation expansion in India. *Environmental Research Letters*, Accepted Manuscript, August 2024.

[13] Fiona Burlig and Louis Preonas. Out of the Darkness and Into the Light? Development Effects of Rural Electrification. *Energy Institue at Haas Working Paper*, page 54, 2016.

[14] Kenneth Lee, Edward Miguel, and Catherine Wolfram. Does Household Electrification Supercharge Economic Development? *Journal of Economic Perspectives*, 34(1):122–144, February 2020.

[15] Sudatta Ray and Hemant K. Pullabhotla. The changing impact of rural electrification on Indian agriculture. *Nature Communications*, 14(1):6780, October 2023.

[16] Abhijit Banerjee and Lakshmi Iyer. History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *THE AMERICAN ECONOMIC REVIEW*, 95(4):24, 2005.

[17] Alice Z. Xu. Segregation and the Spatial Externalities of Inequality: A Theory of Interdependence and Public Goods in Cities. *American Political Science Review*, 118(3):1431–1448, August 2024.

[18] G. S. Ghurye. *Scheduled tribes*.

[19] Virginius Xaxa. Transformation of tribes in india: Terms of discourse. 34(24):1519–1524. Publisher: Economic and Political Weekly.

[20] T.K. Oommen. Scheduled castes, scheduled tribes, and the nation: Situating g. s. ghurye. 60(2):228–244. Publisher: [Sage Publications, Ltd., Sage Publications, Inc.].

[21] Virginius Xaxa. Tribes as indigenous people of india. 34(51):3589–3595. Publisher: Economic and Political Weekly.

[22] Vinita Damodaran. Environment, ethnicity and history in chotanagpur, india, 1850-1970. 3(3):273–298.

[23] Vinita Damodaran. Looking beyond the apocalypse: Environmental crisis, colonial environmentalism and eastern india's tribal communities. In *The Environmental Apocalypse*. Routledge. Num Pages: 18.

[24] Vinita Damodaran and Felix Padel. Investment-induced displacement in central india: A study in extractive capitalism. 38(2):396–411. Publisher: Duke University Press.

[25] Yoko Kijima. Caste and Tribe Inequality: Evidence from India, 1983–1999. *Economic Development and Cultural Change*, 54(2), 2006.

[26] Abhijit Banerjee and Rohini Somanathan. The political economy of public goods: Some evidence from India. *Journal of Development Economics*, 82(2), 2007.

[27] Abhijit Banerjee, Lakshmi Iyer, and Rohini Somanathan. Public Action for Public Goods. *NBER Wroking Paper Series*, Working Paper 12911:40, February 2007.

[28] Alberto Alesina and Ekaterina Zhuravskaya. Segregation and the Quality of Government in a Cross Section of Countries. *American Economic Review*, 101(5):1872–1911, August 2011.

[29] Naveen Bharathi, Deepak Malghan, Sumit Mishra, and Andaleeb Rahman. Status inequality and public goods. *World Development*, 176, 2024.

[30] Sam Asher, Kritarth Jha, Paul Novosad, Anjali Adukia, and Brandon Tan. Residential Segregation and Unequal Access to Local Public Services in India: Evidence from 1.5m Neighborhoods. *Working Paper*, 2024.

[31] Sam Asher, Tobias Lunt, Ryu Matsuura, and Paul Novosad. Development research at high geographic resolution: an analysis of night-lights, firms, and poverty in india using the shrug open data platform. *The World Bank Economic Review*, 35(4), 2021.

[32] Comptroller and Auditor General of India (CAG). Performance audit of accelerated rural water supply programme (ARWSP). 2008.

[33] Government of India Minstry of Drinking Water {and} Sanitation. National rural drinking water programme (NRDWP)_guidelines.

[34] TNM Staff. K'taka minister wants 1-km deep borewell dug for water, clueless that it could worsen the problem.

[35] Minor Irrigation Statistics Division. Manual For Data Collection in the Census of Minor Irrigation Schemes 2013-14. Technical report, Ministry of Water Resources, Government of India, New Delhi, 2017.

[36] Veena Srinivasan, N R Lakshmikantha, G Manjunatha, and Ganesh Shinde. Chasing the water table- The impact of groundwater depletion on rural drinking water supply in peninsular India. *Earth ArXiv Preprint*, 2023.

[37] Rural Domestic Water Supply in India: Progress and Issues. In *Drinking Water Security in Rural India: Dynamics, Influencing Factors, and Improvement Strategy*, Water Resources Development and Management. Springer Nature Singapore, Singapore, 2022.

[38] Divya Balasubramaniam, Santanu Chatterjee, and David B. Mustard. Got Water? Social Divisions and Access to Public Goods in Rural India. *Economica*, 81(321), 2014.

[39] Ashish Kumar Sedai. Who benefits from piped water in the house? Empirical evidence from a gendered analysis in India. ADBI Working Paper Series, 2021.

[40] Pallavi Choudhuri and Sonalde Desai. Lack of access to clean fuel and piped water and children's educational outcomes in rural India. *World Development*, 145, 2021.

[41] Jyotsna Jalan and Martin Ravallion. Does piped water reduce diarrhea for children in rural India? *Journal of Econometrics*, 2003.

[42] Tushaar Shah, O. P. Singh, and Aditi Mukherji. Some aspects of South Asia's groundwater irrigation economy: analyses from a survey in India, Pakistan, Nepal Terai and Bangladesh. *Hydrogeology Journal*, 14(3):286–309, March 2006.

[43] Tushaar Shah and Shilp Verma. Co-Management of Electricity and Groundwater: An Assessment of Gujarat's Jyotirgram Scheme. page 8, 2008.

[44] Abhijit Mukherjee, editor. *Groundwater of South Asia*. Springer Hydrogeology. Springer Singapore, Singapore, 2018.

[45] Soumya Balasubramanya and David Stifel. Viewpoint: Water, agriculture & poverty in an era of climate change: Why do we know so little? *Food Policy*, 93:101905, May 2020.

[46] Reena Badiani and Katrina Jessoe. The Impact of Electricity Subsidies on Groundwater Extraction and Agricultural Production. *Working Paper, University of California at Davis*, page 41, 2011.

[47] Reena Badiani and Katrina Jessoe. Electricity Prices, Groundwater, and Agriculture: The Environmental and Agricultural Impacts of Electricity Subsidies in India. In Wolfram Schlenker, editor, *Agricultural Productivity and Producer Behavior*. University of Chicago Press, 2019.

[48] Nicholas Ryan and Anant Sudarshan. Rationing the Commons. *Journal of Political Economy*, Forthcoming(Mimeo), June 2020.

[49] World Bank. Deep wells and prudence: Towards pragmatic action for addressing groundwater overexploitation in india.

[50] Government of India Ministry of Jal Shakti. Operational guidelines for the implementation of jal jeevan mission JJM.

[51] Robyn Horan, Pawan S. Wable, Veena Srinivasan, Helen E. Baron, Virginie J. D. Keller, Kaushal K. Garg, Nathan Rickards, Mike Simpson, Helen A. Houghton-Carr, and H. Gwyn Rees. Modelling Small-Scale Storage Interventions in Semi-Arid India at the Basin Scale. *Sustainability*, 13(11):6129, May 2021.

[52] Saul Arlosoroff, Gerhard Tschannerl, David Grey, William Journey, Andrew Karp, Otto Langenegger, and Robert Roche. Community Water Supply: The Handpump Option. Technical report, The World Bank, Washington DC, May 1987.

[53] R. Allan Freeze and John A. Cherry. *Groundwater*. Prentice-Hall, Englewood Cliffs, N.J, 1979.

[54] Michael Green and Amit Kapoor. An assessment of aspiration districts programme.

[55] Government of India Minstry of Drinking Water {and} Sanitation. National rural drinking water programme (NRDWP)_guidelines.

[56] Comptroller and Auditor General of India (CAG). Performance audit of the national rurual drinking water programme (nrdwp).

[57] C Dimiceli, M Carroll, R Sohlberg, D Kim, M Kelly, and J Townshend. Mod44b modis/terra vegetation continuous fields yearly l3 global 250 m sin grid v006 [data set]. *NASA EOSDIS Land Process*, 2015.

[58] Government of India Census of India. Census of India 2011: Metadata, 2011.

[59] Central Groundwater Board CGWB. Blockwise Groundwater Reource and Utilization 2011, 2011.

Appendix

Main results - Tables

In this section, we present full tabular results for the coefficient plots presented in the Results section.

Table A1: ST villages are far less likely to use pumped sources for drinking water

	% Pumped water		% Handpump		% Uncovered Wells		% Covered Wells		% Surface		% Other	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ST village	-0.058***	-0.046***	-0.005	-0.006	0.021***	0.015***	0.002*	0.002*	0.039***	0.034***	0.002***	0.002***
	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.001)	(0.001)
SC village	0.006***	0.005**	0.017***	0.017***	-0.017***	-0.017***	-0.001***	-0.001***	-0.004***	-0.004***	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Non-ST/SC population	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total households	0.000***	0.000***	0.000	0.000	0.000	0.000	0.000	0.000	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Road	0.032***	0.028***	0.005	0.005	-0.010***	-0.008***	-0.002***	-0.002***	-0.024***	-0.022***	-0.001**	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Village area	0.000*	0.000	0.000***	0.000***	0.000**	0.000**	0.000	0.000	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to district HQ	-0.001***	-0.001***	0.000	0.000	0.001***	0.000***	0.000	0.000	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
% Electricity		0.113***		-0.007		-0.054***		0.000		-0.045***		-0.007***
		(0.005)		(0.005)		(0.004)		(0.001)		(0.003)		(0.001)
Sub-district fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dependent variable mean	0.269	0.269	0.433	0.433	0.214	0.214	0.020	0.020	0.051	0.051	0.013	0.013
Observations	233,622	233,622	233,622	233,622	233,622	233,622	233,622	233,622	233,622	233,622	233,622	233,622
R^2	0.481	0.486	0.442	0.442	0.337	0.339	0.080	0.080	0.277	0.281	0.242	0.242
Within R^2	0.025	0.034	0.005	0.005	0.005	0.007	0.000	0.000	0.015	0.020	0.000	0.001

Notes: Linear regression of the proportion of households using drinking water source on the presence of ST village binary estimated using equation 1 and dataset constructed using SHRUG [31] and Census 2011 data. We present results for pumped water, handpump, and uncovered wells in the main results but this table includes the remaining sources of drinking water as outcomes as well. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. Standard errors are clustered at the district level. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Table A2: ST villages are less likely to use electricity for lighting

	% Electricity	% Kerosene	Night-time luminosity (logged)
	(1)	(2)	(3)
ST village	-0.107***	0.092***	-0.112***
	(0.004)	(0.004)	(0.008)
SC village	0.005***	-0.006***	0.034***
	(0.002)	(0.002)	(0.004)
Non-ST/SC population	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Total households	0.000*	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Road	0.040***	-0.029***	0.044***
	(0.003)	(0.003)	(0.005)
Village area	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Distance to district HQ	-0.002***	0.001***	-0.005***
	(0.000)	(0.000)	(0.000)
Sub-district fixed effects	Yes	Yes	Yes
Dependent variable mean	0.524	0.453	1.377
Observations	233,622	233,622	233,889
R^2	0.586	0.564	0.673
Within R^2	0.058	0.042	0.109

Notes: Outcomes for Models 1 and 2 are proportion of households using electricity and kerosene as a lighting source respectively, both of which are taken from Census 2011 data. Outcome in Model 3 is the night light luminosity of villages, averaged over the period 2010-2012, as captured by the Defense Meteorological Satellite Program (DMSP) Operation Linescan System (OLS). All models are linear regressions of the outcome on the presence of ST village binary estimated using equation 1 and dataset constructed from SHRUG and Census data [31]. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. Clustered (Sub-district) standard-errors in parentheses. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Table A3: ST villages are far less likely to use pumped sources for drinking water

	% Pumped water		% Electricity
	(1)	(2)	(3)
Hard rock	0.004	0.007	-0.033***
	(0.008)	(0.008)	(0.006)
ST village	-0.034***	-0.025***	-0.107***
	(0.006)	(0.006)	(0.007)
SC village	0.001	0.002	-0.017***
	(0.004)	(0.004)	(0.003)
Non-ST/SC population	0.000*	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Total households	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Village area	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Road	0.018***	0.015**	0.036***
	(0.006)	(0.006)	(0.005)
Distance to district HQ	-0.001***	0.000***	-0.001***
	(0.000)	(0.000)	(0.000)
% Electricity		0.089***	
		(0.012)	
Sub-district fixed effects	Yes	Yes	Yes
Dependent variable mean	0.245	0.245	0.298
Observations	44,634	44,634	44,634
R^2	0.134	0.136	0.511
Within R^2	0.007	0.009	0.080

Notes: Outcomes in Models 1 and 2 are proportion of households in a village using pumped water as primary source of drinking water. Outcome in Model 3 is the proportion of households using electricity as primary source of lighting. All models are linear regressions of the outcome on the presence of ST village binary estimated using equation 8 and dataset constructed from SHRUG and Census data. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. Clustered (sub-district) standard-errors in parentheses. Significance codes: ***: 0.01, **: 0.05, *: 0.1

We decompose the disparity in source of drinking water and lighting at various geographic levels to reveal where the level at which disparity occurs to the greatest extent. We use methods suggested by [30], which are detailed in the Methods section.

Table A4: ST villages are less likely to use pumped water for drinking water across all adminstrative levels

	% Pumped water					
	OLS			Logit		
	No FE	District FE	Sub-district FE	No FE	District FE	Sub-district FE
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.160***			-1.601***		
	(0.006)			(0.035)		
ST village	-0.023***	-0.026***	-0.025***	-0.152***	-0.166***	-0.167***
	(0.005)	(0.008)	(0.006)	(0.026)	(0.049)	(0.038)
SC village	-0.001	0.002	0.002	-0.003	0.009	0.009
	(0.004)	(0.006)	(0.004)	(0.021)	(0.034)	(0.025)
Hard rock	0.038***	0.007	0.007	0.206***	0.039	0.045
	(0.003)	(0.009)	(0.008)	(0.019)	(0.049)	(0.047)
% Electricity	0.152***	0.105***	0.089***	0.812***	0.548***	0.480***
	(0.007)	(0.017)	(0.012)	(0.036)	(0.088)	(0.063)
Non-ST/SC population	0.000	0.000	0.000***	0.000	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total households	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Village area	0.000	0.000*	0.000***	0.000	0.000*	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Road	0.011***	0.018**	0.015**	0.071***	0.106**	0.093**
	(0.004)	(0.008)	(0.006)	(0.024)	(0.046)	(0.040)
Distance to district HQ	0.000	0.000	0.000***	0.000	-0.001	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
District fixed effects		Yes			Yes	
Sub-district fixed effects			Yes			Yes
Dependent variable mean	0.245	0.245	0.245	0.245	0.245	0.245
Observations	44,634	44,634	44,634	44,634	44,634	44,634
R^2	0.024	0.065	0.136			
Within R^2		0.012	0.009			

Notes: Outcomes is proportion of households using pumped water as the primary source of drinking water. Model 1 has no fixed effects, Model 2, includes district fixed effects, Models 3 includes sub-district fixed effects. Models 1-3 are linear regressions estimated using equation 8 and dataset constructed using SHRUG [31] and Census 2011 data. Models 4-6 are corresponding logistic regressions. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. Standard-errors in parentheses, clustered at the same level as fixed effects. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Table A5: ST villages are less likely electricity for lighting across all adminstrative levels

	% Electricity					
	OLS			Logit		
	No FE	District FE	Sub-district FE	No FE	District FE	Sub-district FE
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.388***			-0.352		
	(0.004)			(2.333)		
ST village	-0.134***	-0.112***	-0.107***	-0.951	-0.937***	-0.920***
	(0.003)	(0.014)	(0.007)	(2.166)	(0.124)	(0.065)
SC village	0.005**	-0.005	-0.017***	0.012	-0.046	-0.086***
	(0.003)	(0.007)	(0.003)	(1.489)	(0.033)	(0.017)
Hard rock	-0.089***	-0.040***	-0.033***	-0.458	-0.242***	-0.219***
	(0.002)	(0.011)	(0.006)	(1.408)	(0.071)	(0.038)
Non-ST/SC population	0.000***	0.000***	0.000***	0.001	0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)
Total households	0.000***	0.000	0.000***	-0.001	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.010)	(0.000)	(0.000)
Village area	0.000***	0.000***	0.000***	-0.001	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)
Road	0.045***	0.041***	0.036***	0.250	0.249***	0.235***
	(0.003)	(0.007)	(0.005)	(1.878)	(0.058)	(0.033)
Distance to district HQ	-0.001***	-0.001***	-0.001***	-0.006	-0.006***	-0.008***
	(0.000)	(0.000)	(0.000)	(0.021)	(0.001)	(0.001)
District fixed effects		Yes			Yes	
Sub-district fixed effects			Yes			Yes
Dependent variable mean	0.298	0.298	0.298	0.298	0.298	0.298
Observations	44,634	44,634	44,634	44,634	44,634	44,632
R^2	0.276	0.436	0.511			
Within R^2		0.115	0.080			

Notes: Outcomes is proportion of households electricity as primary source of lighting. Model 1 has no fixed effects, Model 2, includes district fixed effects, Models 3 includes sub-district fixed effects. Models 1-3 are linear regressions estimated using equation 8 and dataset constructed using SHRUG [31] and Census 2011 data. Models 4-6 are corresponding logistic regressions. ST village is a binary indicator for $\geq$ 80% ST population and SC village indicates $\geq$ 20% SC population. Standard-errors in parentheses, clustered at the same as level fixed effects. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Robustness checks

Using proportions of ST and SC population instead of ST and SC village binaries

Table A6: OLS regression of access to drinking water on proportion of ST and SC population

	% Pumped water		% Handpump		% Uncovered Wells	
	(1)	(2)	(3)	(4)	(5)	(6)
ST population share	-0.001***	-0.001***	0.000	0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SC population share	0.000	0.000	0.001***	0.001***	0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Non-ST/SC population	0.000**	0.000**	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total households	0.000***	0.000***	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Road	0.032***	0.028***	0.005*	0.006*	-0.010***	-0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Village area	0.000*	0.000	0.000***	0.000***	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to district HQ	-0.001***	-0.001***	0.000	0.000	0.001***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
% Electricity		0.109***		-0.004		-0.054***
		(0.005)		(0.005)		(0.004)
Sub-district fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Dependent variable mean	0.269	0.269	0.433	0.433	0.214	0.214
Observations	233,622	233,622	233,622	233,622	233,622	233,622
R^2	0.481	0.486	0.442	0.442	0.338	0.339
Within R^2	0.026	0.035	0.005	0.005	0.005	0.008

Notes: Outcomes are proportion of households reporting using respective water structure as main source of drinking water in Census 2011. All models are linear regressions of the outcome on the village population shares of ST and SC (as proportions of total population) using data from SHRUG and Census 2011 [31]. Clustered (Sub-district) standard-errors in parentheses. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Table A7: OLS regression of access to lighting on proportion of ST and SC population

	Electricity	Kerosene	Night-time luminosity
	(1)	(2)	(3)
ST population share	-0.002***	0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
SC population share	-0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Non-ST/SC population	0.000	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Total households	0.000***	0.000**	0.000***
	(0.000)	(0.000)	(0.000)
Road	0.038***	-0.027***	0.042***
	(0.003)	(0.003)	(0.005)
Village area	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Distance to district HQ	-0.002***	0.001***	-0.005***
	(0.000)	(0.000)	(0.000)
Sub-district fixed effects	Yes	Yes	Yes
Dependent variable mean	0.524	0.453	1.377
Observations	233,622	233,622	233,889
R^2	0.591	0.568	0.676
Within R^2	0.069	0.051	0.115

Notes: Outcomes in Models 1 and 2 are proportion of households using electricity and kerosene as a lighting source respectively, both of which are taken from Census 2011 data. Outcome in Model 3 is the night light luminosity of villages, averaged over the period 2010-2012, as captured by the Defense Meteorological Satellite Program (DMSP) Operation Linescan System (OLS). All models are linear regressions of the outcome on the village population shares of ST and SC (as proportions of total population) using dataset from SHRUG and Census 2011 [31]. Clustered (Sub-district) standard-errors in parentheses. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Using logistic regressions instead of OLS

Table A8: Logistic regression of access to drinking water on ST and SC village binaries

	% Pumped water		% Handpump		% Uncovered Wells	
	(1)	(2)	(3)	(4)	(5)	(6)
ST village	-0.568***	-0.462***	-0.057**	-0.061***	0.111***	0.075***
	(0.038)	(0.037)	(0.024)	(0.023)	(0.021)	(0.021)
SC village	0.028**	0.027*	0.084***	0.085***	-0.138***	-0.134***
	(0.014)	(0.014)	(0.012)	(0.012)	(0.014)	(0.014)
Non-ST/SC population	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total households	0.000***	0.000***	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Road	0.242***	0.210***	0.047***	0.049***	-0.072***	-0.059***
	(0.021)	(0.020)	(0.016)	(0.016)	(0.016)	(0.016)
Village area	0.000**	0.000	0.000***	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Distance to district HQ	-0.006***	-0.005***	-0.001	-0.001	0.004***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
% Electricity		1.010***		-0.035		-0.347***
		(0.035)		(0.026)		(0.027)
Sub-district fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Dependent variable mean	0.269	0.269	0.433	0.433	0.214	0.214
Observations	233,618	233,618	233,600	233,600	233,609	233,609

Notes: Outcomes are logistic regressions of the proportion of households reporting using respective water structure as main source of drinking water in Census 2011 on ST and SC village binaries. Expressed as marginal effects, after controlling for the availability of electricity, households in ST villages are 1.5% less likely to use pumped water than non-SC/ST villages and 1.7% more likely to use uncovered wells. On the other hand, SC villages are 0.1% more likely to use pumped water than non-SC/ST villages and 2.9% less likely to use uncovered wells. Clustered (Sub-district) standard-errors in parentheses. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Table A9: Logistic regression of access to lighting on ST and SC village binaries

	Electricity	Kerosene
	(1)	(2)
ST village	-0.616***	0.508***
	(0.025)	(0.024)
SC village	0.032***	-0.036***
	(0.010)	(0.010)
Non-ST/SC population	0.000***	0.000***
	(0.000)	(0.000)
Total households	0.000	0.000***
	(0.000)	(0.000)
Road	0.203***	-0.135***
	(0.015)	(0.014)
Village area	0.000***	0.000***
	(0.000)	(0.000)
Distance to district HQ	-0.010***	0.008***
	(0.000)	(0.000)
Sub-district fixed effects	Yes	Yes
Dependent variable mean	0.524	0.453
Observations	233,620	233,620

Notes: Outcomes are logistic regressions of the proportion of households reporting using electricity and kerosene as main source of lighting in Census 2011 on ST and SC village binaries. Expressed as marginal effects, households in ST villages are 12% less likely to electricity than non-SC/ST villages and 11.1% more likely to kerosene. On the other hand, SC villages are 0.7% more likely to use electricity than non-SC/ST villages and 0.9% less likely to use kerosene. Clustered (Sub-district) standard-errors in parentheses. Significance codes: ***: 0.01, **: 0.05, *: 0.1

Additional results

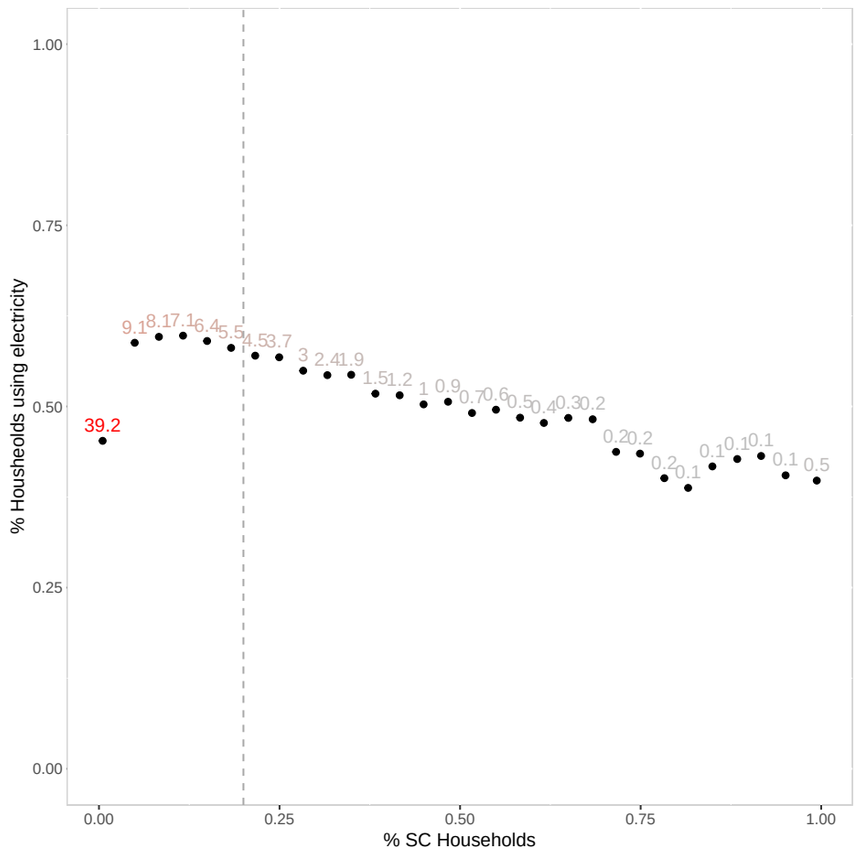


Figure A1: Binned scatter-plot of village-level proportion of SC population and households using electricity. Data labels indicate the proportion of data-points represented by the point.

Table A10: Balance table of villages across granite and alluvial hydrogeologies in Mayurbhanj

		Alluvial	Granite	p-value
1	Households per village	152.76	161.99	0.2
2	Village area	130.78	216.3	0.00
3	ST village (%)	24.47	45.89	0.00
4	SC village (%)	20.89	7.79	0.00
5	All-weather/pucca roads	0.94	0.92	0.23
6	%Electricity	33.01	17.97	0.00
7	%Kerosene	65.81	81.13	0.00
8	%Pumped water	29.31	19.57	0.00
9	%Handpumps	31.55	42.02	0.00
10	%Uncovered wells	33.58	28.92	0.00
11	%Village area covered	0.92	0.97	0.00
12	Total villages	474	1848	NA

Notes: In this table we compare alluvial dominated villages with granite dominated villages in one district of Odisha, Mayurbhanj. Means of the various characteristics as well as p-values of t-tests are provided. Mayurbhanj has a total of 3,672 villages and is one of the largest districts of the state, constituting approximately 7% of the total state area.

Table A11: Range of district transmissitivity values in Odisha

Rock type	Range	Classification
Alluvial	2 to 8189	Non-hard
Bedded Sandstone/Quartzite	17.4 to 114.7	Non-hard
BHJ/BCQ	5 to 50	Hard
Charnockite	0.6 to 3691.7	Hard
Fruchsite/Sericite Quarzite	7.8 to 3691.7	Hard
Gneiss	0 to 3691.7	Non-hard
Granite	0.6 to 3691.7	Hard
Khondalite	0.3 to 70.2	Hard
Laterite	2.8 to 490.2	Non-hard
Meta basic/Meta volcanic	0.71 to 81	Hard
Phyllite /Mica Schist	2 to 60	Non-hard
Quartzite	0.6 to 60	Hard
Sandstone	0.9 to 6	Non-hard
Sandstone/Conglomerate/Shale	0.5 to 305	Non-hard
Shale/Phyllite	2 to 3	Non-hard