Got Water? Social Divisions and Access to Public Goods in Rural India[*]

Divya Balasubramaniam[♣]
University of Georgia

Santanu Chatterjee[♦]
University of Georgia

David B. Mustard[♠]
University of Georgia

August 28, 2009

Abstract

We examine whether different social divisions help explain the variation in tap water access across India. Using data for 436 rural districts from the 2001 Census of India, we find that communities that are heterogeneous in terms of Hindu caste have less access to tap water than correspondingly homogeneous communities. By contrast, religiously fragmented communities have more access to tap water than correspondingly homogeneous communities. Therefore, heterogeneity *within* and *across* religions may work in opposite directions for access to public goods. Consequently, the many studies that use aggregate measures of social fragmentation may obscure important information regarding the design of public policy related to public goods.

Keywords: Public goods, Social fragmentation, Water, Public policy, India
JEL Classification: H4, O2

[*]We are grateful to Rohini Somanathan for providing the Hindu Caste data. This paper has benefited from presentations at the University of Georgia and the Southern Economic Association annual meeting in Washington, DC. Balasubramaniam acknowledges financial assistance provided by the Dean's Award at the University of Georgia, which enabled the purchase of the Indian Census and Meteorological data.

[♣]Department of Economics, Terry College of Business, University of Georgia, Athens, GA 30602. Email: divya@uga.edu

[♦]Department of Economics, Terry College of Business, University of Georgia, Athens, GA 30602. Email: schatt@terry.uga.edu

[♠] Corresponding author: Department of Economics, Terry College of Business, University of Georgia, Athens, GA 30602. Email: mustard@terry.uga.edu

1. Introduction

Water is essential for economic activity and an important determinant of health outcomes and living standards. In most developed countries, water services are provided through a pricing mechanism. By contrast, in many developing countries, the government acts as the sole provider for water services. India represents a prime example of a fast-growing developing country where water is a "public" good, i.e., whose provision by the government is designed to provide universal and free access to its pre-dominantly poor and rural population. The data, however, reveal wide variation in access to drinking (tap) water across rural India, with some districts having no access at all. The central objective of this paper therefore is to understand the extent to which social and economic factors in rural India determine access to water services.

Our study focuses on the roles played by caste and religion in determining access to publicly provided water, which are important in India's social setting. To this extent, our paper is related to recent work by Banerjee and Somanathan (2007), who examine how access to public goods across rural India changed between 1971 and 1991. However, our study differs from existing work in several very important dimensions.

Most studies use an *aggregate* index of social fragmentation, which is made up of several socio-economic characteristics such as ethnicity, race, language, religion, caste, etc. We argue that the aggregate fragmentation index is not useful for designing public policy, as it does not convey information on the magnitude and the direction of the individual characteristics that comprise the index. Different types of social divisions may have disparate effects on the provision of (and access to) public goods, and also call for different policy measures. In that sense, our study represents a new approach whereby we use disaggregated measures of social fragmentation: do measures of diversity based on caste, religion, and political preferences move

in the same or different directions when determining access to water services in rural India? What are their individual effects on access to drinking water? In essence, we examine how fragmentation *within* a religion (e.g. caste system among Hindus) compares with fragmentation *across* religions (Hindus, Muslims, Christians, etc) in determining access to public goods.

We use data from the 2001 Census of India for 436 rural districts to test for the effect of caste and religious heterogeneity on tap water access in rural India.[1] We use three measures of tap water access: the share of households in a district having access to (i) total tap water, (ii) tap water within a residence, and (iii) tap water outside a residence. Compared to previous studies that use either aggregated measures of fragmentation or public goods, we use household data and disaggregated measures of both public good access and social fragmentation. In this respect, our approach provides a more granular view of the issue at hand. The main result of this paper is that caste heterogeneity reduces and religious heterogeneity increases tap water access in rural India.

Districts that are more heterogeneous in terms of caste have significantly lower access to both total tap water and within-residence tap water.[2] A 10 percentage-point increase in caste-based heterogeneity in a district reduces a household's probability of total tap water access by 3.9 percentage points. On the other hand, districts that are more heterogeneous in terms of religion have significantly higher access to total tap water and within-residence tap water: a 10 percentage-point increase in religion-based heterogeneity in a district increases a household's probability of total tap water access by 3.1 percentage points.[3]

[1] The analysis is restricted to rural areas because the caste classification, which is one of the primary variables of interest, is available only for rural India.

[2] The heterogeneity index reflects the mean within-group affinity for a public good. The higher the share of an individual's own group in the population, the higher is the probability that he or she will have access to public goods.

[3] This result is consistent with Alesina et al. (2003), who find that countries with higher degrees of religious heterogeneity also tend to have a higher quality of infrastructure services.

Another contribution of this paper is that disaggregating the standard measure of social fragmentation helps resolve a puzzle reported by Banerjee and Somanathan (2007), who find that while their aggregate social fragmentation index reduces access to public goods in 1971, this effect is statistically insignificant in 1991. They interpret this result as indicating that between 1971 and 1991, social and religious groups in India were able to mobilize themselves politically, which diminished the importance of social fragmentation over time. However, our results provide an alternative interpretation: when employing an aggregate measure of social fragmentation, the effects of the individual components may offset each other. Indeed, even 10 years beyond Banerjee and Somanathan's period of study (1971-91), we find that social fragmentation does matter, with its different components moving in different directions. This result is only evident when one examines the relative contribution of *each* component of the fragmentation measure. Our results indicate that public policy must be designed to target different aspects of social divisions, rather than the one-size-fits-all approach often adopted.

Finally, our results also highlight the extent to which social factors influence the *source* of drinking water: for access to tap water *within* the residence, the role of caste heterogeneity is crucial; by contrast, for access to tap water *outside* the residence, what really matters is the concentration of different caste groups.

The rest of the paper is organized as follows. Section 2 explains the institutional set up and water crisis in India. Section 3 discusses the nature and depth of social divisions in India and the possible mechanisms through which they affect access to drinking water. Section 4 describes the empirical specification, while section 5 characterizes the data and discusses some econometric issues such as endogeneity and selective migration. Section 6 analyzes the results and the

underlying intuition, while Section 7 concludes with a brief discussion of how private provision of water services might overcome the barriers imposed by social divisions.

2. India's Institutional Setup and Water Crisis

2.1. Water Institutions

Since independence in 1947, the provision of water services in India has predominantly been under the control of the government. India follows a top-down approach of water management, where the central government has a monopoly in providing water with some limited degree of decentralization (Saleth, 2005). Each state is responsible for the delivery of water within its state boundary. The state may, however, vest this responsibility to the Panchayat Raj Institutions (PRI) in the rural areas or to the municipalities (urban local governments) in urban areas.[4] The delivery of water in each state is overseen by the Department of Public Health Engineering, Rural Development Engineering, or a Water Board. Even though the states generally plan, design, and execute water supply schemes, final decision-making and approvals are dictated by the central government, which coordinates investment in the water sector though it's Five Year Plans.[5] Several institutions within the central government, such as the Central Water Commission, closely coordinate and regulate surface water use in various sectors like industry, irrigation, drinking water, etc. They act as mediators in inter-state water disputes as well. The Central Ground Water Board (CGWB), created in the year 1972, is responsible for ground water management.

[4] As per the Indian Constitution, the PRI have certain powers and authority at the village level to devise plans that ensure economic development and social justice.

[5]The Five Year Plan, designed and implemented by the Planning Commission of India, is a statement of the economic development targets and objectives for five years from the date of implementation.

At the national level, the Ministry of Rural Development delivers water to rural areas through the Department of Drinking Water Supply. The department controls issues related to water policies, provides funds for investment in water projects, and manages and assists states in water supply to rural areas. In the urban sector, the Ministry of Urban Development (MoUD) is responsible for water supply. Responsibilities include planning, setting standards, monitoring and support state programs by providing financial and technical expertise. The Ministry of Housing and Urban Poverty alleviation is responsible for urban water supply in the union territories. However, there is no autonomous body to regulate the water sector in India (Government of India, 2002).

2.2. The Water Crisis in India

Several recent studies indicate a looming water crisis in India, with demand far outstripping supply. India has experienced enormous growth in its agriculture, industry, and service sectors since independence. Particularly, economic liberalization since the early 1990's and the sustained economic growth India has experienced in the last two decades has resulted in a dramatic increase in the demand for water from all sectors of the economy: consumers, industry, agriculture, and services. However, investments in the water supply infrastructure have failed to keep pace with the rapid growth elsewhere in the economy, thereby increasing the scarcity of a critical resource (Bajpai, 2007). Moreover, the lack of private provision in the water sector, especially in the rural areas, has further exacerbated the problem (Swiss Agency for Development and Cooperation, 2007).

Due to rapid urbanization, industrialization, and a lack of enforceable environmental laws, most Indian rivers are polluted and non-potable (Ramachandran, 2006). The World Bank estimates that unsafe water is responsible for 21 percent of communicable diseases in developing

countries (World Bank, 1993). Although the Indian government made huge investments in the water infrastructure during the 1960's, since then it has allocated very little for operations and maintenance, which has resulted in an aging infrastructure. Haarmeyer and Mody (1997) review how governments could draw on private capital to address inefficiencies in the water sector. They argue that water sectors controlled by governments in developing countries are financially and operationally weak, because they collect revenues that cover only 35 percent of the total cost of water production. They highlight the need for a more efficient institutional arrangement. Because the Indian government lacks the ability to supply adequate water and no law restricts the amount of ground water extracted by a land owner (Saleth, 2005), many middle-income people in urban areas extract private ground water. Over-extraction of ground water has depleted water tables and magnified the crisis. Inter-state legal disputes about water sharing have further worsened this crisis. Another issue is global warming, which several environmentalists believe causes erratic climatic conditions and aggravates the water crisis (Brooks, 2007). All these trends have spurred significant concerns about the current provision of water services.

3. Soci-religious Stratification and its Effects on Access to Drinking Water

The data reveal a stark story about water access for rural India: according to the 2001 Census of India, the average share of households in a rural district that had access to either inside or outside tap water ranged from 3 percent in the eastern state of Orissa to 83 percent in the northern state of Himachal Pradesh. The numbers for tap water access *within* a residence are even worse: from 0 percent in Orissa to 27 percent in the western state of Maharashtra. Not only does access to government-provided water services vary widely across the country, no rural district has 100 percent access to tap water (See Table 1). Despite its geo-tropical position and being endowed with one of the world's largest river networks, the supply of water in India has

been woefully inadequate (Sengupta, 2006). The majority of India's population lives in rural areas and depends heavily on publicly provided water. As such, a sustained water shortage can amplify economic hardship and intensify competition among social groups for a scarce but essential public commodity. The wide variation in the share of rural households with access to tap water is a strong reflection of this point. Many economists believe that the government's inadequacy in providing an adequate and uniform supply of water to India's rapidly growing economy poses one of the biggest threats to its potential for progress and prosperity.

Even though one expects economic factors such as poverty and inequality to play a pivotal role in determining access to public goods, these factors are intricately linked to historically persistent social divisions (Banerjee et al., 2005). India has had a long history of invasions (from Mongolia and Persia) and external occupation (French, Dutch, Portuguese, and British), which ended with Independence from British colonial rule in 1947. Consequently, Indian society is deeply fragmented along social and religious lines, and these divisions play a pivotal role in both politics and the allocation of scarce public resources more than six decades after independence.

The predominant form of social stratification in India is the caste system, which has deep historic roots in the majority Hindu religion. Historically, the objective to create caste divisions was to facilitate the identification of social groups based on their skill levels. However, over time, certain groups gained power over others, forming the basis for social discrimination. Over time, stratification based on caste became the determinant of an individual's position in the social hierarchy. Even though discrimination based on caste is illegal in India, there is still a sharp "social" difference between "high" and "low" castes. According to the Census of India,

there are more than 180 caste groups within the Hindu religion: *Brahmans* belong to the highest caste, while *Scheduled Castes* and *Scheduled Tribes* are at the bottom of the social hierarchy.[6]

Another dimension of social division in India is religion. Even though the Indian Constitution provides for a secular state with equal tolerance for all religions, more than 80 percent of the population is Hindu, while the principal minorities include Muslims, Christians, Sikhs and Buddhists (Census of India, 2001). Religious conflicts, often based on economic discrimination between Hindus and Muslims, and more recently between Hindus and Christians, have been historically pervasive in India. India's multi-party democracy, with regional and caste-based political parties often playing important roles in national politics, and along with high levels of poverty and inequality, underscore the nature and depth of its social fragmentation.

Why might social (caste) and religious divisions determine access to public goods such as water? The answer lies in the nature of social stratification and the competition it creates for scarce public goods. Several authors, starting with Easterly and Levine (1997) and Alesina et al. (1999) highlight the need for coordination within communities to gain access to public goods. In our context, therefore, social divisions based on caste and religion can, in theory, either impede or facilitate this coordination mechanism. On the one hand, social fragmentation may lead to explicit or implicit competition for scarce public goods and, over time, may lead the dominant castes and religious groups to appropriate much of the benefits of public goods. Water, being essential to economic activity, is therefore at the center of this competition. On the other hand, strong secular and legal institutions could also lead to more tolerance and cooperation amongst social or religious groups over time, so that communities may benefit from this cooperation in

[6]Brahmans are at the highest end of the caste spectrum and have been historically associated with being priests, teachers, and philosophers. At the other end of the spectrum lie the Scheduled Caste and Scheduled Tribes. These are terms the Government of India uses to classify the poorest and most disadvantaged communities in India; Scheduled Caste refers to the *Dalit* community and Scheduled Tribe to the tribal communities or *adivasis*.

the form of higher access to public goods. Therefore, it is not clear how different dimensions of social divisions (*within* a religion and *across* religions) affect access to public goods. Our paper sheds new light on this important public policy issue and highlights the underlying mechanisms through which different social divisions affect access to drinking water in rural India.[7]

4. Empirical Specification

The central focus of this paper is to identify the determinants of access to tap water in rural India. The specification we test is given by the following cross-section regression:

$$Y_{i,k} = \alpha_1 HI_{i,k} + \alpha_2 S_{i,k} + \alpha_3 R_{i,k} + \alpha_4 E_{i,k} + \alpha_5 P_{i,k} + \alpha_6 X_{i,k} + \alpha_7 M_k + \varepsilon_{i,k}$$

$Y_{i,k}$ is the share of households with access to tap water (within and outside the residence) in district *i* of state *k*. $HI_{i,k}$ is the Hindu caste and/or religion *homogeneity* index in a given district of a state. The index measures the probability that two distinct individuals picked randomly from the population in a given district belong to the same (a) Hindu caste, or (b) religion. This is similar to a Herfindahl index and is constructed by calculating

$$HI_j = \sum_{j=1}^{n} s_j^2$$

where s_j is the population share of the *j*-th caste or religious group. Therefore, a positive coefficient estimate indicates that higher social homogeneity (based on caste or religion) increases tap water access, or conversely, higher social heterogeneity decreases tap water access, since the corresponding heterogeneity index is simply equal to 1-homogeneity index. $S_{i,k}$ represents a set of social variables, including population shares of the Hindu caste groups (Brahman, Scheduled Caste and Scheduled Tribes). $R_{i,k}$ contains the shares of Muslims,

[7] In this context, our paper is related to research on the importance of social heterogeneity for public good provision; see Dayton-Johnson (2000), Miguel and Gugerty (2005), and Khwaja (2009) for some recent contributions.

Christians and other religious minorities in the sample. $E_{i,k}$ represents a set of economic variables, including the land Gini coefficient (to capture inequality) and bank deposits per capita. $P_{i,k}$ is a set of election outcomes that include the share of total votes cast that were received by the winning party and a political heterogeneity index.[8] $X_{i,k}$ is a vector of geographical characteristics that include average annual rainfall, average temperatures, terrain, a dummy for coastal areas, and controls for population density including the average village population, average number of villages, and household size. M_k is a set of state fixed-effects, and $\varepsilon_{i,k}$ is a district-specific shock.

The empirical specification we adopt is quite standard in the social heterogeneity-public goods literature; see Alesina et al., (1999). Though there is conclusive evidence on the effects of ethnic heterogeneity with respect to public goods, the interpretation of these effects is complicated by the nature of the aggregation mechanism that links individual preferences to community outcomes. Vigdor (2004) shows how the theory of altruistic behavior can be used to aggregate individual decisions (to contribute to public goods) to community outcomes. We therefore assume that the effects of group heterogeneity on access to public goods (tap water, in our case) are derived from the altruistic behavior of individuals in each community (district).[9]

Another important point of the model specification is the inclusion of individual group shares, with individual coefficient estimates for each group, and a single coefficient for the homogeneity (or heterogeneity) index. The homogeneity index reflects the mean within-group affinity for a public good. The higher the share of an individual's own group in the population, the higher is the probability that he or she will contribute to a given set of public goods. If the

[8] The political heterogeneity index estimates the probability that two individuals randomly drawn from a population will belong to different political parties.
[9] Banerjee and Somanathan (2007) and Rushton (2008) also adopt a similar procedure.

model fails to control for the group shares, it then imposes an implicit behavioral restriction that all groups have the same propensity to contribute to public goods, regardless of the composition (caste or religious) of the community. Consequently, the Hindu caste shares enter the specification in two important ways. First, following Vigdor (2002, 2004), the shares of Hindu caste groups appear as a linear term with a separate coefficient for each caste group.[10] Second, they appear as squared terms in the caste homogeneity index, with a single coefficient. We include individual shares in the specification to capture the effect of the presence of other caste groups on the access to tap water for individuals in a specific group. The homogeneity index captures how the number of groups affects access to water, assuming that the within-group affinity is equal across all the groups in the district.[11] When both the index and the individual shares are included in the same regression, we cannot make a ceteris paribus argument, since when the share of a group changes, the homogeneity index will also change simultaneously. We follow the same procedure to analyze and interpret the effects of religious heterogeneity.

We estimate the model with ordinary least squares, and discuss related econometric issues such as endogeneity, in detail in the next section.

5. Data

We use district-level data from the 2001 Census of India, the latest year for which data are available. Our dataset includes 436 of the 593 rural districts in India, since data for some of the explanatory variables were not available for districts in several states. Table 1 reports the

[10]Because we use 180 caste groups, the inclusion of each caste group separately in the model may be complex. Therefore, we aggregate these 180 groups into three sub-groups: Brahmans, Scheduled Castes, and Scheduled Tribes. Each of these three sub-groups appears linearly in the specification, which is consistent with the literature.

[11] Ideally, one could include the shares and their squared terms with separate coefficients in the regression. The squared group shares capture the within-group-affinity across groups and a coefficient estimate for each squared group share allows for affinity to vary across caste groups. We do not make this assumption because of the large number of caste groups in our sample and including each of the 180 groups and its squared term is very complex.

summary statistics for the variables we use and, as mentioned in the introduction, reveals some striking patterns with respect to access to tap water across rural India.

5.1. Tap Water Access

The share of households with access to tap water is the main dependent variable for our analysis, and is obtained from the Houses, Households and Amenities Section of the 2001 Census of India. We use three measures of tap water access in our study: (i) total tap water access, (ii) within-residence tap water access, and (iii) outside-residence tap water access.[12] Table 2 shows that the mean share of a household's total tap water access ranges from 3 percent in the eastern state of Orissa to 83 percent in the northern state of Himachal Pradesh; the mean share of a household's within-residence tap water access ranges from nearly 0 percent in the state of Orissa to 27 percent in the western state of Maharashtra. The mean share of a household's outside-residence tap water access ranges between 5 percent and 58 percent across the sample. These numbers not only reflect large variation across Indian states, but also document that none of these states have 100 percent access to tap water in their rural districts. The segment of the population not served by tap water uses hand pumps, wells, rivers or other water sources to meet their daily water needs. Outside-residence tap water access implies substantial costs borne by the households, including travel time to the water source and waiting time to get access to water.

5.2. Caste and Religion

Detailed caste data in India were last enumerated under the British Colonial regime in 1931. After independence, caste-based data collection was discontinued to prevent discrimination and, from 1951 onwards, the Indian government has collected data based on three broad categories: Scheduled Castes, Scheduled Tribes and Others. In calculating the caste homogeneity index, we

[12] Total access to tap water includes both within and outside the residence access. Outside-residence tap water access refers to a household traveling 100 meters or more in rural areas to fetch drinking water.

use the methodology described in Banerjee and Somanathan (2007).[13] Because the caste data are from the 1931 Census, and a significant Muslim population immigrated to Pakistan after Independence in 1947, they adjust the increase in the proportion of Hindus after 1931 by scaling up the numbers in each caste group based on the Hindu share in the current census.[14] We similarly scale up the caste figures by the share of Hindu population in 2001 and also adjust for newly created districts between 1991 and 2001. In all, we have 180 caste groups within the Hindu religion in our sample.

One important point of departure from the Banerjee-Somanathan study is in the construction of the caste homogeneity index. Their study combines 185 Hindu caste groups with six non-Hindu religions in the construction of a "socio-religious" heterogeneity index, thus assuming that other religious (non-Hindu) groups are internally homogeneous. We do not make any assumptions about the caste structure in other religions, but instead restrict our analysis to only the Hindu castes. Further, we also do not combine other forms of heterogeneity such as language and religion in constructing the caste index because in India there are many languages with several sub-dialects and it is very difficult to group people or communities by language.[15]

[13] We are grateful to Rohini Somanathan for sharing the caste data used in the Banerjee-Somanathan study. The 1931 Census had a very large list of caste groups for each British province and princely state, by district. Over time, even though state boundaries were redrawn, district boundaries remained intact. After independence, a few districts were created by subdividing old ones. For these new districts, Banerjee and Somanathan weigh the caste data by the area of the new district that was created from the original districts. Since the number of caste groups is very large, they restrict the number to the Hindu caste that constitutes more than one percent of the population of each state or province in 1931. This approach yields 185 Hindu caste groups. Since Hindus are a majority, this restriction is reasonable. We use data on 180 out of the 185 caste groups as some states were not included in the sample due to lack of data on other crucial explanatory variables.

[14] This assumes that over time, all Hindu castes grew at a similar rate. Since this paper focuses on rural areas, the method is a reasonable approximation, since the percentage of rural to urban migration has been relatively slow in India (Haub and Sharma, 2006).

[15] Easterly and Levine (1997) use measures of ethno-linguistic heterogeneity constructed from the former Soviet Union in 1960. The data, however, were based on linguistic classification rather than on race or color. One problem with this measure is that it may hide other aspects of ethnicity; see Alesina et al. (2003). For example, if two ethnic groups speak the same language but have different customs and beliefs, then classification based purely on language combines these two different ethnic groups in one category, which in turn may generate measurement error.

Data on the various religious groups (Hindus, Muslims, Christians, and others) are from the religion tables of the Census of India 2001. The religion index may have potential problems because a person can potentially hide his or her religion to avoid oppression. Individuals may change from one religion to another while it is historically less likely that people change from one caste to another (Alesina, et al. 2003). We address this issue in detail in section 3.2.

Table 1 shows that the Hindu caste homogeneity index ranges between 0 and 0.32, with a mean of 0.04, which implies that rural districts are highly heterogeneous in terms of caste. The religion homogeneity index ranges between 0.37and 0.99 with a mean of 0.80. Since 85 percent of the population is Hindu, the sample mean of 0.80 implies that districts are highly religiously homogeneous.

5.3. Economic Controls

Private wealth, an indicator of economic status, might be an important determinant of access to water. To this end, per-capita bank deposits across rural districts in 2001 are obtained from the Reserve Bank of India database. The number and area of operated land holdings by different sizes (measured in hectares) are obtained from the 2001 Agricultural Census of India. We calculate the land Gini coefficient using these data to proxy for land inequality across districts. We assign zero land holdings to agricultural laborers. Because there are no data on ownership land holdings, the use of operated land holdings may be less than a perfect measure for land distribution. However, one defense of this variable is that since Independence most land on average has been owner-cultivated (Banerjee and Somanathan, 2007). Table 1 shows that the land Gini ranges between 0.41 and 0.86 with a sample mean of 0.71, implying that rural districts have a high degree land inequality.

5.4. Political Controls

In India, political power is crucial in determining access to water across districts. Politicians are concerned about the number of votes they will receive in the next election based on the satisfaction of the public on the various public goods provided during their last term in office. Data on general elections for 1999 are from the Election Commission of India (1999) website.[16] We use two political variables, namely (i) the vote share of the winning party, and (ii) a political fragmentation index. The index is calculated using:

$$p = 1 - \sum_{i=1}^{n} v_i^2$$

where v_i is the vote share of the *i*-th party. The index *p* lies between 0 and 1, where 1 represents complete political heterogeneity and 0 represents political homogeneity. The political fragmentation index ranges between 0.14 and 0.80, with a mean of 0.59, thereby implying a high degree of political heterogeneity across rural districts (Table 1).

5.5. Geography Controls

Since average rainfall and temperature affects access to water, we also control for these measures. Data on average annual rainfall and average temperatures are from two sources: (i) The Indian meteorological department (IMD) and (ii) rainfall and average temperature maps released by the Maps of India website.[17] The share of wastelands in each district determines the level of water access as well. The shares of land that are steep, barren, and sandy were obtained

[16] We use 1999 because general elections are conducted once in five years and 1999 is the closest year to the analysis period. The data on general elections are available for the 543 parliamentary constituencies in India. We use a mapping method from Banerjee and Somanathan (2007) that makes the data consistent at the district level. Specifically, the data are mapped by visually comparing the number of districts that go to each parliamentary constituency. We further compare the number of constituencies that go to each district by assigning weights by visual inspection using maps of districts and maps of parliamentary constituencies.

[17] We map the available data from 109 weather stations in India on to districts in the following manner. First, we assign average rainfall and average temperature values to each district based on its proximity to each weather station. Second, for a few states in the northeastern region where no data is available, we use the median value calculated using the maps that contain the average annual rainfall and average temperatures.

from the Wasteland Atlas of India (Ministry of Rural Development) for 2003.[18] Finally, we also include a coastline dummy.

5.6. Other Population Controls

Average household size, the number of villages in each district, and the average village population from the 2001 Census of India captures the population density in a district, since population composition may affect access to tap water.

5.7. Correlation between Dependent and Control Variables

Table 3 shows the correlation between the Hindu caste homogeneity index, religion homogeneity index, and each of the three dependent variables. Of the three dependent variables, the within-residence tap water share has the largest positive correlation with the Hindu caste homogeneity index (0.295). The total tap water share and outside residence tap water share have a very high correlation of 0.91, which implies that most of the tap water access for the given sample is from outside the residence. The Brahman share is positively correlated with tap water share access, because Brahmans are historically considered the upper-most caste group in India and we expect them to have greater access to public goods, including tap water. The religion homogeneity index does not bear any consistent correlations with the dependent variables.

Table 4 shows the correlation coefficients between the dependent variables and economic and political variables. It reveals consistent signs for each of the variables. One expects a negative sign for the land Gini, which implies that the higher is land inequality, the lower is the access to tap water. Similarly, the higher is political heterogeneity, the lower is the access to tap water. Other economic variables such as per-capita bank deposits are also important because

[18] According to the Ministry of Rural Development of India sandy areas have stabilized accumulation of sand, in coastal, riverine, or inland areas, and steepy lands are steep sloping wasteland areas.

wealthier districts may have more public funding and hence better access to tap water. However, the magnitude of these correlations is small.

5.8. Endogeneity

A potential econometric issue is endogeneity, either through omitted variables, simultaneity, or measurement error. Hence, it is important to determine if the effect of Hindu caste and religion heterogeneity on tap water access is robust to correcting for these issues. The caste-based classifications in India created during pre-historic times are determined at the time of birth, and mobility across castes is prohibited by social norms. Therefore, the Hindu caste homogeneity index is not likely to be endogenous. Moreover, the Hindu caste variable is based on 1931 data, and therefore, is unlikely to be influenced by contemporaneous changes in districts. A more serious issue is the endogeneity of the religion homogeneity index. Data on religious conversion rates over time may solve the issue, but are not available for India. However, the religion data between 1961 and 2001 show that the population proportions across religions have been very stable.[19] Therefore, it is also unlikely that the religious homogeneity index is endogenous.

5.9. Selective Migration

There is also the problem of selective migration because people may prefer to migrate to districts that have higher access to tap water. This can influence the caste group shares and the size of the village. Haub and Sharma (2006) show that rural-urban migration has been very low in India, especially when compared to countries in Latin America. Urban migration in India rose from 11 percent in 1901 to only 28 percent in 2001. Most Indians live their entire lives in rural areas. However, the lack of disaggregated data on net migration hinders a richer specification for

[19]The Census of India (2001) shows that the proportion of Hindus (including both rural and urban areas) in 1961 was 84.4 percent, while in 2001 it was 81.4 percent; proportion of Muslims in 1961 was 10.2 percent and in 2001 it was 12.4 percent; proportion of Christians in 1961 was 2.4 percent and in 2001 it was 2.3 percent. Other religious groups grew at a similar rate between 1961 and 2001.

our analysis. To test for selective migration, we exclude the population density variables from the model specification to check for the robustness of the results. We also include a wide range of other controls. Since we restrict the analysis to only rural areas, this minimizes the effects of urbanization on water access. All regressions include state fixed-effects, unless otherwise noted, to account for unobserved characteristics across states, and the standard errors are clustered to account for within-state unobserved variation.

Finally, we compare the two homogeneity indices with Alesina et al. (2003). They find a small but positive correlation between their measures of ethnic and religion fragmentation (0.142). In our sample, there is also a small positive correlation between the Hindu caste homogeneity index and the religion homogeneity index (0.201).

6. Results and Discussion

6.1. The Hindu Caste Homogeneity Index

Table 5 shows the results from the ordinary least squares regressions where the share of total tap water is the dependent variable. Since the test for homosckedacticity in all the regressions was rejected, we cluster errors by state and calculate Huber-White standard errors. Column (1) in Table 5 shows the coefficient estimates on the social variables, column (2) adds the religion, economic and political controls, and column (3) contains the entire set of controls in the regression. The coefficient estimate of the Hindu caste homogeneity index is positive and statistically significant at the 10 percent level in all three columns: a household's probability of access to (total) tap water increases by 3.9 percentage points when the household's own caste share of the population increases by 10 percentage points (column 1). The magnitude of this coefficient estimate changes only by 0.3 percentage points between columns (2) and (3). The

estimated coefficient of Scheduled Tribes is statistically significant at the 5 percent level in all the columns. The point estimates imply that districts with a high concentration of Scheduled Tribes have relatively lower access to total tap water.

Inclusion of the entire set of controls improves the model's fit, increasing the overall R^2 from 0.009 to 0.051. Adding the geography, population, and wasteland controls only marginally changes the magnitude on the coefficient estimates of the social variables. The geography controls, especially the coefficient estimates of average rainfall and coastline controls, are negative and statistically significant at the 5 percent level. This is counter-intuitive because we expect that higher rainfall and proximity to a coast will increase access to water. One plausible reason for this may be the rainfall and temperature calculations. Since the calculations are based on visual mapping of weather stations to districts, the values may not be accurate. The religion, population density, economic, and political controls are not statistically significant.

Table 6 reports the results of the ordinary least squares regression where within-residence tap water access share is the dependent variable. Even after controlling for all covariates, the Hindu caste homogeneity index remains positive and statistically significant at the 5 percent level: a 10 percentage-point increase in a households' own caste share of the population increases the household's probability of within-residence tap water access by 2.5 percentage points. The differences in magnitude are small across the three columns. The vote share of the winning political party and the political fragmentation index are both positive and statistically significant at the 5 percent level. The religion, economic, and population controls are not statistically significant.

Table 7 shows results using outside-residence tap water access as the dependent variable. The most interesting result here is for the Scheduled Tribes, which are among the most

disadvantaged groups. The coefficient estimate is negative and statistically significant at the 5 percent level in all three columns. This is interesting because given the position of Scheduled Tribes in the socio-economic scale in India, their main source of drinking water is likely to be outside the residence. The coefficient estimate on the Hindu caste homogeneity index is positive but statistically insignificant. Therefore, in the case of access to outside-residence tap water, the concentration of different groups matter more than the number of groups. The religion, economic, political, and population controls are statistically insignificant. Inclusion of the entire set of controls improves the model's fit, increasing the overall R^2 from 0.02 to 0.15.

6.2 The Religion Homogeneity Index

Table 8 provides the regression results for the three types of tap water access, but with the religion homogeneity index as the main explanatory variable. All three regressions include the entire set of controls discussed earlier. The results show that more homogenous districts in terms of religion have lower access to tap water. In other words, controlling for other factors, households in districts that are more fragmented on religious lines have higher access to tap water than those in more religiously homogeneous districts. Column (1) shows the results for total tap water access: a 10 percentage point increase in a household's own religion group in the population decreases its probability of access to total tap water by 3.1 percentage points. The estimated coefficient is statistically significant at the 5 percent level. The results for within-residence tap water access, in column (2), are also similar and statistically significant at the 1 percent level. In column (3), for outside-residence tap water access, the effect of the religion homogeneity index is negative but is statistically insignificant. In all three columns, the share of Scheduled Tribes is negative and statistically significant at the 10 percent and 5 percent level in columns (1) and (3), respectively.

The results for tap water access for different religious groups are interesting. In all three regressions the relative tap water access for Muslims is lower compared to Hindus and is statistically significant at the 5 percent level for both total and within-residence access to tap water and at the 10 percent level for outside-residence tap water access. These results are plausible because Muslims are a prominent minority religion group and there are more frequent social conflicts between Hindus and Muslims than between any other religious groups. The more intriguing results are for the Christians. Their relative access to all three types of tap water is lower and statistically significant at the 5 and 10 percent levels in columns (1) and (2), respectively. In general, social tensions between Hindus and Christians are less frequent and therefore we do not expect any form of access problems for Christians. But the results indicate that they too have lower relative access to tap water. On the other hand, Sikhs, who represent another minority religion (breakaway from Hindus), have higher access to both total and within-residence tap water (statistically significant at the 10 and 5 percent levels, respectively).

Finally, we also include both the Hindu caste homogeneity and religion homogeneity indices with the shares of castes and religion groups in the same regression (table not reported).[20] The estimated coefficients on all of the social and religion variables are similar to the ones previously reported and consistent.

6.3. Robustness Checks

We perform a number of sensitivity checks to ensure that the results from using the two measures of social divisions are robust. Table 9 reports the coefficient estimate on the Hindu caste homogeneity index from a range of sensitivity tests. Specification (1) shows the baseline estimates. Specification (2) shows the coefficient estimates when the population density

[20] Results are available upon request.

variables are excluded from the regressions. The results are robust to dropping the population density variables. Particularly, the coefficient estimate of Hindu caste homogeneity index is positive and is statistically significant at the 10 percent level. We also conduct many other specification checks. Since the correlation between the two political control variables is high (-0.9), in some specifications we include either the political index or the vote share of the winning party. These specifications seldom change the results of the variable of interest. Further, both the political variables were also excluded from the specification due to the possibility of endogeneity. But the results remain intact.

Because the population of a village influences caste composition in a given district, we include an interaction term between the Hindu caste homogeneity index and the average village population, to check if the magnitude of the average village population affects the partial effect of the Hindu caste homogeneity index. However, there is no evidence from this sample that the magnitude of the average village population influences the effect of caste heterogeneity on tap water access.

Table 10 reports the results of an analogous sensitivity analysis for the religion homogeneity index. All the results, excluding the interaction between the religion index and the average village population, are robust to specification changes. The interaction term shows evidence that the magnitude of the average village population influences the partial effect of religion heterogeneity on tap water access.

6.4. Caste versus Religion: the Role of History and Institutions

One of the striking findings of our empirical exercise is that while social divisions based on the Hindu caste system reduces access to tap water in rural India, those based on religion seem to improve access. A natural question at this point is: why do the results go in opposite directions?

A possible explanation may lie in the role played by history and institutions in determining the impact of social divisions on the coordination mechanism needed by a community for gaining access to public goods.

As discussed earlier, the caste system in India has been historically pervasive and has been the source of segregation and intolerance in society from ancient times. Indeed, the practice of "untouchability," whereby people belonging to "lower" castes were prohibited from interacting socially with those from the "higher" castes, provided the foundation for this segregation and lack of tolerance. For example, a person belonging to a lower caste was prohibited from entering the house of a person who was higher up in the caste hierarchy. These historically pervasive social divisions create social norms over time, which not only tend to be very persistent, but also are critical in determining economic outcomes; see Ray (1998, chapter 5). Our results on the effects of caste heterogeneity therefore indicate that, even with the caste system being constitutionally illegal in India, its historical barriers have prevented the cooperation necessary within communities to get access to publicly provided drinking water.

Religious diversity, on the other hand, plays a very different role in affecting the coordination mechanism needed for access to public goods. India has a long history of trade with Persia and Europe, as well as external conquests from the very same regions. These also exposed the country very early on to the world's predominant religions, such as Islam and Christianity. The Mughal Empire ruled India for almost 400 years, and was followed by British colonization for 200 more (which also coincided with some areas being under French, Dutch, and Portuguese occupation). The existing social order and the need for external commerce perhaps created a degree of tolerance among religions in India that eventually became historically persistent. The secular institutions guaranteed by the Indian Constitution (Freedom of Religion being a

Fundamental Right), further strengthened the tolerance and cooperation among religions since Independence. Our result on the religion heterogeneity index underscores this point.

7. Conclusions

We examine whether different aspects of social divisions in India help explain the wide variation in access to tap water across rural India. In contrast to most studies, which use *aggregate* measures of social fragmentation that are comprised of several socio-economic characteristics such as ethnicity, race, language, religion, and caste, we employ disaggregate measures. Consequently, our approach better allows for individual measures of diversity to have heterogeneous effects on outcomes.

The empirical analysis suggests that communities that are heterogeneous in terms of caste within the majority Hindu religion are likely to have lower access to tap water than correspondingly homogeneous communities. By contrast, communities that are fragmented across religions are likely to have more access to tap water than correspondingly homogeneous communities. In essence, even though Hindus are a large majority among religious groups in India, representing more than 80 percent of the population, our results indicate that religious diversity fosters better access to public goods, perhaps because secular institutions generate greater tolerance between people across religions. These are interesting results, since they indicate that though both heterogeneity *within* and *across* religions matter for access to public goods, but they may work in opposite directions. Consequently, studies that use an aggregated measure of social fragmentation by combining many characteristics of social divisions are unlikely to reveal reliable information regarding its impact. Our results also indicate that Scheduled Tribes and minority religious groups such as Christians and Muslims have relatively lower access to tap water in rural India compared to Hindus. Therefore, while caste-based and

religious heterogeneity is important for understanding public goods access, the existence of minority groups in the economy (in terms of economic or religious classifications) also matter. Finally, the interaction between social factors and the source of tap water is also important: while caste-based fragmentation is crucial for tap water access within the residence, the concentration of caste groups matters for tap water outside the residence.

These results point to the need for public policy reform in the water sector in India. Given that certain types of social divisions like the caste system creates barriers for adequate public provision of drinking water, should public policy encourage private participation in the water sector? Two recent examples strengthen this view. Davis et al. (2008) survey 800 poor households in the southern city of Hyderbabad. A large majority reported inadequacies in government provided water and sanitation. Interestingly, their regression analysis suggests that even if faced with non-concessional market rates of financing, these households would prefer to pay for private investment in water and sewer connections. These results underscore the vital role micro-financing can play in overcoming social barriers. Another example comes from the region of Tirupur in southern India. A recent public-private partnership has ensured the supply of drinking water for 4-6 hours each day for 80,000 households, compared to getting water every alternate day of the week before the partnership. More interestingly, 100% of the residents (mostly poor) now pay for the water (Mulford, 2006).

We end with a caveat. Because we use district-level data, it is not possible for us to identify the underlying mechanism that drives the opposite signs for the group heterogeneity results (caste and religion). The problem is that, for local public goods like water, the measurement of social heterogeneity and water access would ideally occur at a smaller kilometer grid because each district may contain multiple communities and therefore investments in one community

may not benefit others (Jackson, 2007). The lack of data at a more disaggregated level for rural India precludes a solution for this problem. However, in the wake of the current water crisis in India, these results provide insights into the role played by two important sources of social fragmentation in India, namely caste and religion, and will, in turn, direct future research to analyze the underlying mechanism that drives these opposing effects of heterogeneity.

References

Agricultural Census of India. 2001. Department of Agriculture & Cooperation, Agricultural Census Division, Ministry of Agriculture, India.

Alesina, Alberto, Baqir, Reza and Easterly, William. 1999. "Public goods and ethnic divisions." *Quarterly Journal of Economics* 114 (4): 1243–1284.

Alesina, Alberto, Devleeschauwer, Arnaud, Easterly, William, Kurlat, Sergio and Wacziarg, Romain. 2003. "Fractionalization." *Journal of Economic Growth*, 8, 155–194.

Bajpai, Nirupam. 2007. "India's Growing water crisis." *Hindu Business Line*, 12, January.

Banerjee, Abhijit, Iyer, Lakshmi and Somanathan, Rohini. 2005. "History, Social Divisions, and Public Goods in Rural India." *Journal of the European Economic Association*, 3(2-3), 639–647.

Banerjee, Abhijit and Somanathan, Rohini. 2007. "The Political Economy of Public Goods: Some Evidence from India." *Journal of Development Economics*, 82, 287-314.

Brooks, Nina. 2007. "Imminent water Crisis in India." http://www.arlingtoninstitute.org/wbp/global-water-crisis/606#

Census of India. 2001. Office of the Registrar General & Census Commissioner, India.

Davis, Jenna, Gary White, Said Damodaran, and Rich Thorsten. 2008. “Improving access to water supply and sanitation in urban India: microfinance for water and sanitation infrastructure development.” *Water Science and Technology*, 58, 887-891.

Dayton-Johnson, J. 2000. “Determinants of collective action on the local commons: a model with evidence from Mexico.” *Journal of Development Economics*, 62, 181-208.

Easterly, William, and Levine, Ross. 1997. “Africa’s Growth Tragedy: Policies and Ethnic Divisions,” *The Quarterly Journal of Economics*, 112(4), 1203–1250.

Election Commission of India. 1999. Statistical Reports of Lok Sabha Elections. http://www.eci.gov.in/StatisticalReports/ElectionStatistics.asp

Government of India. 2002. India- Assessment-Water and Sanitation. Planning Commission, Government of India.

Haarmeyer, David and Mody, Ashoka. 1997. “Private Capital in Water and Sanitation.” *Finance and Development.*

Haub, Carl and Sharma, O.P. 2006. “India’s Population Reality: Reconciling Change and Tradition,” *Population Bulletin*. 61(3), 1-20.

Jackson, Ken. 2007. “Why Does Diversity Affect Public Good Provision? An Empirical Analysis of Water Provision in Africa.” Dissertation Chapter, University of British Columbia.

Khwaja, A. 2009. “Can good projects succeed in bad communities?” *Journal of Public Economics*, 93, 899-916.

Miguel, Edward, and Gugerty, Mary Kay. 2005. “Ethnic Diversity, Social Sanctions and Public Goods in Kenya.” *Journal of Public Economics*, 89, 2325–2368.

Mulford, David. 2006. “Public-Private Partnership can help solve India’s water and sanitation challenges.” *The Hindu,* March 22 issue.

Ramachandran, Asha. 2006. “Any Hope for India’s Water Woes.” India/Kerala News. http://www.greentreks.org/woodchuckcafe/features/feature_indiaswater.asp

Ray, Debraj. 1998. *Development Economics*. Princeton University Press, Princeton, NJ.

Rushton, Michael. 2008. “A Note on the Use and Misuse of the Racial Diversity Index.” *Policy Studies Journal,* Forthcoming.

Saleth, Rathinasamy Maria. 2005. “Water institutions in India: Structure, performance, and change.” In Gopalakrishnan, C.; Biswas, A. K.; Tortajada, C. eds., *Water Institutions: Policies, Performance and Prospects*. New York, NY. pp: 47-80.

Sengupta, Somini. 2006. “In Teeming India, Water Crisis Means Dry Pipes and Foul Sludge.” *New York Times*, 29, September.

Swiss Agency for Development and Cooperation (SDC). 2007. “Retention and sustainable use of water.” http://www.sdcindia.in/en/Home/Activities_in_2007/Water_NRM

Vigdor, Jacob. L. 2002. “Interpreting Ethnic Fragmentation Effects.” *Economic Letters*, 75, 271-276.

Vigdor, Jacob. L. 2004. “Community Composition and Collective Action: Analyzing Initial Mail Response to the 2000 Census.” *The Review of Economics and Statistics*, 86, 303–312.

World Bank. 1993. World Development report: “Investing in health”. New York: Oxford University Press.

World Bank Report. 2005. “India’s Water Economy: Bracing for a Turbulent Future.”

Wasteland Atlas of India. 2003. Ministry of Rural Development, Department of Land Resources, Government of India. http://dolr.nic.in/fwastecatg.htm

Table 1: Summary Statistics

Variables	Number of Observations	Mean	Std. Dev.	Min.	Max.
Dependent Variables					
Total tap share	436	.27	.24	.01	.91
Total Tap Share (Within)	436	.11	.11	0	.57
Total Tap Share (Outside)	436	.17	.18	0	.80
Independent Variables					
Caste Variables					
Hindu Caste Index (Homogeneity)	436	.04	.05	0	.32
Brahman	436	.04	.04	0	.27
Scheduled Caste	436	.11	.09	0	.46
Scheduled Tribe	436	.04	.06	0	.53
Religion Variables					
Religion Index (Homogeneity)	436	.80	.16	.37	.99
Hindus	436	.85	.18	.04	.99
Muslims	436	.09	.12	0	.79
Christians	436	.02	.06	0	.47
Sikhs	436	.03	.15	0	.94
Buddhists	436	.01	.04	0	.59
Jains	436	.002	.004	0	.04
Others	436	.003	.016	0	.19
Not Stated	436	.001	.001	0	.01
Economic Variables					
Land Gini	436	.71	.09	.41	.86
Per Capita Bank Deposits	436	.02	.02	.002	.23
Political Variables					
Political Index	436	.59	.09	.14	.80
Vote Share of Winning Party	436	.48	.08	.27	.71
Population Variables					
Household Size	436	5.38	.85	4	8
Number of Villages (‘000s)	436	1.13	.821	.041	10.54
Avg. Village Population (‘000s)	436	1.98	3.13	.115	26.79
Geography Variables					
Avg. Rainfall (meters)	436	.98	.68	.07	5.88
Avg. Temperature (Celsius)	436	25.44	2.89	14.47	29.88
Coastline	436	.13	.34	0	1
Sandy	436	.04	.11	0	.97
Barren	436	.04	.07	0	.50
Steepy	436	.01	.03	0	.49

Table 2: Mean Share of Households with Access to Tap Water

State Name	No. of Districts	Total	Within	Outside
Andhra Pradesh	22	0.40	0.13	0.27
		(0.18)	(0.08)	(0.13)
Assam	23	0.06	0.01	0.05
		(0.06)	(0.01)	(0.05)
Chattisgarh	16	0.05	0.02	0.03
		(0.02)	(0.01)	(0.02)
Gujarat	25	0.47	0.25	0.22
		(0.23)	(0.15)	(0.12)
Haryana	19	0.38	0.12	0.26
		(0.14)	(0.06)	(0.11)
Himachal Pradesh	12	0.83	0.25	0.58
		(0.06)	(0.09)	(0.11)
Karnataka	27	0.48	0.11	0.37
		(0.17)	(0.04)	(0.15)
Kerala	14	0.14	0.08	0.06
		(0.08)	(0.04)	(0.05)
Madhya Pradesh	45	0.11	0.05	0.06
		(0.09)	(0.05)	(0.04)
Maharashtra	30	0.45	0.27	0.18
		(0.16)	(0.12)	(0.07)
Orissa	29	0.03	0.01	0.02
		(0.01)	(0.01)	(0.01)
Punjab	17	0.17	0.11	0.06
		(0.08)	(0.04)	(0.05)
Rajasthan	32	0.21	0.13	0.08
		(0.13)	(0.09)	(0.06)
Tamilnadu	28	0.6	0.07	0.53
		(0.13)	(0.04)	(0.11)
Uttarkhand	11	0.67	0.19	0.49
		(0.17)	(0.09)	(0.20)
Uttar Pradesh	70	0.15	0.09	0.06
		(0.10)	(0.08)	(0.04)
West Bengal	16	0.08	0.02	0.06
		(0.07)	(0.02)	(0.05)
Total	436	0.27	0.11	0.17
		(0.24)	(0.10)	(0.18)

Source: Census of India, 2001. Standard deviations in parentheses.

Table 3: Correlations between Tap Water Access and Hindu Caste Variables

	Tap share (total)	Tap share (within)	Tap share (outside)	Hindu Caste Index	Brahman	SC	ST	Religious Index
Tap share (total)	1							
Tap share (within)	0.71	1						
Tap share (outside)	0.92	0.36	1					
Hindu Caste Index	0.28	0.29	0.21	1				
Brahman	0.13	0.13	0.1	0.42	1			
SC	0.01	0.08	-0.03	0.25	0.3	1		
ST	0	0.24	-0.14	0.29	0.06	0.24	1	
Religion Index	0.1	-0.03	0.16	0.2	0.28	0.19	0.14	1

Note: SC- Scheduled Caste; ST- Scheduled Tribe

Table 4: Correlations between Tap Water Access and Economic and Political Variables

	Tap share (total)	Tap share (within)	Tap share (outside)	Land Gini	PCBD	Vote share	Political Index
Tap share (total)	1						
Tap share(within)	0.71	1					
Tap share (outside)	0.92	0.36	1				
Land Gini coefficient	-0.11	-0.1	-0.08	1			
PCBD	0.32	0.17	0.32	-0.17	1		
Vote share	0.15	0.04	0.17	0.08	0.14	1	
Political Index	-0.22	-0.06	-0.25	-0.04	-0.19	-0.91	1

Note: PCBD- Per Capita Bank Deposits.

Table 5: Caste Composition and Total Tap Water Access

Dependent Variable: Total Tap water access share

Effect [Independent Variable]	(1)	(2)	(3)
Social Variables			
Mean within group affinity [Caste Homogeneity Index]	0.385*	0.350*	0.297*
	(1.84)	(1.82)	(1.91)
Share of Brahmans	-0.574	-0.520	-0.339
	(-1.58)	(-1.45)	(-1.19)
Share of Scheduled castes	0.131	0.149	0.073
	(0.79)	(0.93)	(0.57)
Share of Scheduled Tribes	-0.277**	-0.272**	-0.296**
	(-2.32)	(-2.23)	(-2.47)
Religion Variables			
Share of Muslims		0.039	0.087
		(0.25)	(0.63)
Share of Christians		-0.022	-0.041
		(-0.11)	(-0.31)
Share of Sikhs		0.195	0.229
		(0.58)	(0.83)
Economic Variables			
Per Capita Bank Deposits		0.0253	0.207
		(0.04)	(0.31)
Land Gini		-0.0411	-0.043
		(-0.32)	(-0.43)
Political Variables			
Vote share		0.203	0.197
		(1.38)	(1.36)
Political Index		0.046	-0.011
		(0.26)	(-0.08)
Geography Variables			
Rainfall (in meters)			-0.042***
			(-2.54)
Temperature (Celsius)			-0.004
			(-1.03)
Coastline			-0.0920**
			(-2.93)
Constant	0.278***	0.170	0.485**
	(21.47)	(0.80)	(2.43)
N	436	436	436
Population Control Variables	NO	NO	YES
Wasteland Variables	NO	NO	YES
State Fixed Effects	YES	YES	YES
Overall R^2	0.009	0.004	0.051

Heteroskedasticity-consistent t-statistics (clustered by state) included in parentheses; *, ** and *** represent 10, 5 and 1% significance level respectively. Columns (2)-(3) include shares of other religion groups.

Table 6: Caste Composition and Within-residence Tap Water Access

Dependent Variable: Tap water access share (Within)

Effect [Independent Variable]	(1)	(2)	(3)
Social Variables			
Mean within group affinity [Caste Homogeneity Index]	0.217	0.266**	0.249**
	(1.22)	(2.41)	(2.07)
Share of Brahmans	-0.332	-0.334**	-0.249
	(-1.41)	(-1.96)	(-1.34)
Share of Scheduled castes	0.055	0.0698	0.047
	(0.67)	(1.37)	(0.84)
Share of Scheduled Tribes	-0.123	-0.126	-0.126
	(-1.26)	(-1.32)	(-1.37)
Religion Variables			
Share of Muslims		0.123	0.134
		(1.16)	(1.26)
Share of Christians		0.091	0.076
		(1.05)	(0.99)
Share of Sikhs		0.265	0.271
		(1.28)	(1.57)
Political Variables			
Vote share		0.205**	0.203**
		(2.31)	(2.27)
Political Index		0.195***	0.173**
		(2.85)	(2.44)
Constant	0.109	-0.172*	-0.024
	(11.43)	(-1.95)	(-0.20)
N	436	436	436
Economic Control Variables	NO	YES	YES
Population Control Variables	NO	NO	YES
Geography Control Variables	NO	NO	YES
Wasteland Control Variables	NO	NO	YES
State Fixed Effects	YES	YES	YES
Overall R^2	0.000	0.000	0.008

Heteroskedasticity- consistent t- statistics (clustered by state) included in parentheses; *, ** and *** represent 10, 5 and 1% significance level respectively. Columns (2)-(3) include shares of all other religion groups.

Table 7: Caste Composition and Outside-residence Tap Water Access

Dependent Variable: Tap water access share (Outside)

Effect [Independent Variable]	(1)	(2)	(3)
Social Variables			
Mean within group affinity [Caste Homogeneity Index]	0.167	0.083	0.048
	(0.71)	(0.40)	(0.28)
Share of Brahmans	-0.242	-0.186	-0.091
	(-0.97)	(-0.69)	(-0.46)
Share of Scheduled castes	0.077	0.079	0.027
	(0.46)	(0.50)	(0.20)
Share of Scheduled Tribes	-0.155**	-0.146**	-0.170**
	(-1.98)	(-2.00)	(-1.99)
Religion Variables			
Share of Muslims		-0.083	-0.047
		(-1.23)	(-1.02)
Share of Christians		-0.114	-0.117
		(-0.90)	(-1.53)
Share of Sikhs		-0.069	-0.041
		(-0.46)	(-0.32)
Constant	0.168***	0.342	0.509***
	(24.85)	(1.59)	(2.63)
N	436	436	436
Economic Control Variables	NO	YES	YES
Political Control Variables	NO	YES	YES
Geography Control Variables	NO	NO	YES
Population Control Variables	NO	NO	YES
Wasteland Variables	NO	NO	YES
State Fixed Effects	YES	YES	YES
Overall R^2	0.021	0.159	0.152

Heteroskedasticity- consistent t- statistics (clustered by state) included in parentheses; *, ** and *** represent 10, 5 and 1% significance level respectively. Columns (2)-(3) include shares of all other religion groups.

Table 8: Religious Composition and Tap Water Access

Using the Religious Homogeneity Index

Effect [Independent Variable]	Tap water Share (Total)	Tap water Share (Within)	Tap water Share (Outside)
Social Variables			
Mean within Group affinity [Religion Homogeneity Index]	-0.309**	-0.260***	-0.049
	(-2.29)	(-3.18)	(-0.60)
Share of Brahmans	-0.157	-0.096	-0.062
	(-0.48)	(-0.54)	(-0.29)
Share of Scheduled castes	0.116	0.082*	0.034
	(0.87)	(1.80)	(0.28)
Share of Scheduled Tribes	-0.223*	-0.065	-0.158**
	(-1.75)	(-0.65)	(-1.99)
Religion Variables			
Share of Muslims	-0.212**	-0.118**	-0.094*
	(-2.24)	(-2.16)	(-1.64)
Share of Christians	-0.427**	-0.250***	-0.178
	(-1.99)	(-2.00)	(-1.57)
Share of Sikhs	0.249*	0.287***	-0.038
	(1.79)	(3.66)	(-0.35)
Constant	0.829***	0.266*	0.563***
	(3.35)	(1.92)	(2.77)
N	436	436	436
Economic Control Variables	YES	YES	YES
Political Control Variables	YES	YES	YES
Geography Control Variables	YES	YES	YES
Population Control Variables	YES	YES	YES
Wasteland Control Variables	YES	YES	YES
State Fixed Effects	YES	YES	YES
Overall R^2	0.028	0.012	0.142

Heteroskedasticity- consistent t- statistics (clustered by state) included in parentheses; *, ** and *** represent 10, 5 and 1% significance level respectively. Column (1)-(3) include shares of all other religion groups (Buddhists, etc.).

Table 9: Robustness Check for the Hindu Caste Homogeneity Index

Specification	Tap water share (Total) (1)	Tap water Share (Within) (2)	Tap water share (Outside) (3)
(1) Baseline (including All controls)	0.297*	0.249**	0.048
	(1.91)	(2.07)	(0.28)
(2) Excluding the population density variables	0.314*	0.249**	0.066
	(1.91)	(2.13)	(0.36)
(3) Excluding Political Index	0.297*	0.234**	0.064
	(1.85)	(1.96)	(0.35)
(4) Excluding Vote share of the winning party	0.287*	0.242**	0.048
	(1.84)	(2.05)	(0.28)
(5) Excluding both the political controls	0.313*	0.238*	0.075
	(1.78)	(1.93)	(0.40)
(6) Using Interaction Terms:			
Average Village Population * Caste Homogeneity Index	0.290	0.094	0.197
	(1.52)	(0.72)	(0.85)
N	436	436	436
State Fixed Effects	YES	YES	YES

Heteroskedasticity- consistent t- statistics (clustered by state) included in parentheses; *, ** and *** represent 10, 5 and 1% significance level respectively.

Table 10: Robustness Check for the Religion Homogeneity Index

Specification	Tap water share (Total) (1)	Tap water Share (Within) (2)	Tap water share (Outside) (3)
(1) Baseline (including all controls)	-.309**	-.260***	-0.048
	(-2.29)	(-3.18)	(-0.60)
(2) Excluding the population density variables	-0.319**	-0.250***	-0.059
	(-2.31)	(-3.04)	(-0.74)
(3) Excluding Political Index	-0.307**	-0.263***	-0.044
	(-2.30)	(-3.15)	(-0.54)
(4) Excluding Vote share of the winning party	-0.311**	-0.261***	-0.049
	(-2.28)	(-3.11)	(-0.60)
(5) Excluding both the political controls	-0.301**	-0.261***	-0.040
	(-2.19)	(-3.13)	(-0.47)
(6) Using Interaction Terms:			
Average Village Population * religion homogeneity Index	-0.399**	-0.314***	-0.084
	(-2.34)	(-3.31)	(-0.79)
N	436	436	436
State Fixed Effects	YES	YES	YES

Heteroskedasticity- consistent t- statistics (clustered by state) included in parentheses; *, ** and *** represent 10, 5 and 1% significance level respectively.