

Ethnic Diversity and Ethnic Discrimination: Explaining Local Public Goods Provision

Alexander Lee*

March 30, 2017

Abstract

Numerous studies have found that ethnic diversity is negatively associated with the provision of local public goods. However, these accounts neglect both the strong role of central institutions in the provision of many “local” public goods and the frequently positive correlation between diversity and the presence of less politically powerful ethnic groups. These factors suggest that existing diversity findings may be explained in some cases by central governments discriminating against areas inhabited by less powerful groups. This hypothesis is tested using data in village-level public goods provision in Northern India, supplemented by data on service provision in Kenyan villages and American cities. While there is evidence that the presence of socially powerful groups is positively associated with service provision, evidence for the diversity hypothesis is weak. The results suggest that failures of public services in diverse areas may reflect larger inequalities within the political system rather than local problems in cooperation.

Word Count: 9996

*Department of Political Science, University of Rochester

1 Introduction

In both wealthy and poor countries, many areas lack access to public services such as roads, schools, health clinics, and electricity. This lack is made particularly troubling by close association between these services, economic growth, and social outcomes. While there are many explanations for poor provision of public services, one in particular has received scholarly attention: the level of ethnic diversity, usually operationalized using a Herfindahl fractionalization index. The negative association between ethnic diversity and public services has been attested for a wide variety of nations, goods, and levels of aggregation (Alesina, Baqir and Easterly, 1999; Banerjee and Somanathan, 2007; Miguel and Gugerty, 2005; Easterly and Levine, 1997; Algan, Hémet and Laitin, 2011) though some recent studies have also questioned the relationship at the local level (Gerring et al., 2015; Gisselquist, Leiderer and Niño-Zarazúa, 2016), and others have argued that the relationship is conditional on the salience of ethnic divisions, which may be a product of national policies, segregation, and between-group economic differences (Miguel, 2004; Singh and vom Hau, 2016; Trounstone, 2015; Baldwin and Huber, 2010). This relationship is usually traced to the inability of people from different ethnic backgrounds to cooperate with each other, a failure variously traced to differences in preferences, lack of a common language, and an inability to sustain cooperative equilibria due to social sanctioning (Habyarimana et al., 2007).

Relative to many other hypotheses in political economy, the diversity-public goods relationship is both empirically well-supported and theoretically grounded. There are, however, two reasons to think that cooperation-based theories do not adequately capture the real life process by which public goods are allocated at many times and in many countries.

Firstly, many local public services are provided through processes that involve national or regional governments much larger than the unit of observation, either through the direct provision, the distribution of money to local governments, or the staffing of facilities constructed by local governments. In the United States, for instance, local schools are constructed by local governments, but often a large portion of the funds used come from the state and federal governments. Indeed, in many nations there are

considerable practical difficulties in identifying “pure” locally provided public goods outside of an experimental setting, and this has led many tests of the diversity hypothesis to focus on allocation processes in which non-local actors play an important role.

Secondly, most diversity theories implicitly assume that groups are politically equal conditional on local group size. However, in many countries, some ethnic groups are more politically and socially powerful than others. In many cases the powerful group are also the national or regional majority group. In France, for instance, white French people are generally more influential than those of Middle Eastern descent. However, the powerful group may also be a small group that for historical reasons is overrepresented in higher education, land ownership, or the bureaucracy. Rwandan Tutsis, for instance are both a minority group and are (today) the most politically powerful group in the country.

Taken together, these implied scope conditions of diversity models suggest an alternative linkage between identity and public goods provision—the tendency of central policy makers to provide goods to areas inhabited by social powerful groups. If policy-makers are disproportionately drawn from these groups, this preference may reflect the well-established tendency of politicians to favor coethnics in service provision (Franck and Rainer, 2012; Hodler and Raschky, 2014; Burgess, Jedwab and Miguel, 2015). However, members of dominant social groups may have superior access to public goods even when their coethnics are not in elected office, due to more subtle processes of social discrimination.

Why, then, have so many studies found a negative diversity effect? Many existing accounts fail to account for the correlation between ethnic power differences and ethnic diversity. In countries or regions where the powerful group is a majority of the population, the local presence of these groups is highly collinear with ethnic fractionalization. In the United States, for instance, cities with high levels of diversity are overwhelmingly those with few whites: fractionalization is negatively correlated with percent white at .86. It is thus often difficult, if not impossible, to separate any disadvantage stemming from diversity per se from the effects of discrimination against non-whites. In many countries, this will mean that more diverse areas have worse public services even if this difference stems from discrimination by central institutions rather than a lack of

cooperation among local people.

To properly test both the discrimination and diversity hypotheses, it is thus necessary to examine a case where 1) the literature has found a negative diversity effect for the types of goods in question, 2) we should expect cooperation problems between ethnic groups, and 3) levels of fractionalization are uncorrelated with the presence of the dominant group. This paper implements this strategy using data on the provision of village-level public goods in the Northern Indian states of Bihar and Uttar Pradesh. Not only are these goods very similar to those examined in other parts of the world, but existing studies strongly support the idea that caste diversity is associated with low public goods provision in India (Banerjee and Somanathan, 2007) and that caste groups have difficulty cooperating with each other (Fehr, Hoff and Kshetramade, 2008). But the most important advantage of Northern India as a case is that social diversity is uncorrelated with the presence of the most politically powerful groups. While members of four high-status caste groups have traditionally been disproportionately powerful in the politics and society of the study region, the small size of these groups mean that their presence is not associated with village-level caste fractionalization ($\rho=-.01$).

The Indian data provides powerful support for the ethnic discrimination hypothesis. While the presence of upper castes has a strong positive association with the most important types public services provision, there is no evidence of a negative association between diversity and public goods, even for goods with little non-local involvement. These results hold even when accounting for variation in the salience of caste identities. Like nearly all studies of the identity-public goods relationship, this study is potentially confounded by geographical and economic differences between areas inhabited by different types and numbers of ethnic groups. To attempt to deal with this issue, a variety of alternative models include controls for spatial differences between villages, village-level wealth, general and intergroup economic inequality, and patronage networks.

These results might be dismissed as a product of some unique aspect of the Indian case. Perhaps India's public goods provision process is too centralized for the diversity hypothesis to hold, but the hypothesis still applies everywhere else? While weighing the relative importance of these two mechanisms on a global scale is impossible, this study provides additional evidence that even in cases cited as evidence for the diversity

approach, the evidence for a diversity effect is limited once we account for discrimination. To do this, the study reanalyses data on public goods provision taken from two influential studies: [Alesina, Baqir and Easterly \(1999\)](#) (on American cities) and [Miguel \(2004\)](#) (on villages in western Kenya). In both studies, ethnic diversity is highly correlated with the presence of the regional or national majority ethnic group. As a consequence, while both ethnic diversity and the presence of the majority ethnic group are associated with public goods provision, the variables are so collinear that both relationships disappear when the two are included in the same model. These results demonstrate the difficulty of separating the power and diversity hypotheses in many cases, and the fact that the discrimination hypothesis is fully consistent with the data even in cases cited as evidence for a diversity effect.

2 Ethnic Politics and Public Goods: Existing Accounts

Governments have the potential to provide a wide variety of goods and services to their citizens that may both have direct welfare benefits and enhance economic growth. Classic examples include education, health care, transportation infrastructure, electricity and water. In political economy discussions, such broadly beneficial goods are often referred to as public goods, even though they do not fit the classic definitions of the concept in economic theory, as many of these goods are in practice both rival and excludable. The underprovision of these goods represents a puzzle for scholars because it is, at least in a societal sense, irrational, as overall welfare could be enhanced by providing them. A variety of theories of this failure have been developed, focusing on factors such as low state capacity, social capital, and self-interested behavior by politically powerful elites. One factor, however, has received a very large share of scholarly attention: The role of ethnic diversity.

2.1 The Diversity Hypothesis

Theoretical accounts of the diversity mechanism differ in details, but share certain broad features. Most importantly, they begin by assuming the public goods provision requires cooperation between a large number of individuals within a locality. This cooperation can take the form either of assembling a minimal winning coalition to impose taxes to pay for the services (Alesina, Baqir and Easterly, 1999), or encouraging a large number of people to voluntarily contribute labor and money in defiance of incentives to free ride (Miguel and Gugerty, 2005). If this cooperation succeeds, the good is provided, while if it fails, it is not.

The central contention of the diversity hypothesis is that cooperation is more likely to occur when the individuals involved are from the same ethnic group. A variety of mechanisms have been proposed to explain this pattern. Coethnics may prefer similar types of goods, may find cooperation easier due to shared languages and social networks, or may be more easily able to sanction free-riding. Since Habyarimana et al. (2007)’s influential contribution, this last mechanism has been the most popular, though all proposed mechanisms give a broadly similar prediction about the relationship between diversity and public goods, and link it to failures in local-level cooperation.

This relationship is supported by a wide variety of analyses, showing diversity to be negatively associated with public goods across a wide variety of units, including countries (Easterly and Levine, 1997), American cities (Alesina, Baqir and Easterly, 1999), African villages (Miguel, 2004), French housing blocks (Algan, Hémet and Laitin, 2011) and Indian districts (Banerjee and Somanathan, 2007). Supplementing the observational work, another literature has shown that non-coethnics are less likely to cooperate with each other or give to each other in laboratory settings (Habyarimana et al., 2007; Whitt and Wilson, 2007).

A well-known critique of these findings is that the ethnic divisions on which they focus are both the product of endogenous social processes and vary in their importance across social contexts. The most extreme form of this argument would hold that the entire diversity-public goods relationship is a product of omitted factors (Wimmer, 2015). In response, scholars have proposed alternative formulations of the diversity argument,

holding that ethnic divisions are only significant when they are made salient by factors such as national policy and history (Miguel, 2004; Singh and vom Hau, 2016), interethnic inequality (Baldwin and Huber, 2010), and residential segregation (Trounstein, 2015).

A small but growing literature has begun to question the validity of the diversity-public goods relationship at the subnational level, where it appears that the association is zero or positive in many contexts (Gerring et al., 2015; Gisselquist, Leiderer and Niño-Zarazúa, 2016). However, these account generally do not question the theoretical mechanism that public goods are cooperatively provided and ethnic differences impede cooperation. Instead, they argue that problems of cooperation at the local level may be ameliorated by the availability of individuals with different sets of skills, or because monitoring and social sanctioning are easier in a smaller communities than in larger ones.

2.2 Criticisms of the Diversity Hypothesis

Despite the status of the diversity hypothesis as “one of the most powerful hypotheses in political economy” (Banerjee, Iyer and Somanathan, 2005)[639], there are several reasons to believe that in many circumstances the ethnic cooperation mechanism does not adequately capture the process by which public goods are produced and consumed in many countries. Two such critiques are detailed below. In common with the literature, these critiques assume that public goods are non-rival and non-excludable; however, a third potential critique, the de facto excludability of many local “public” goods, is discussed in an online appendix.

Centralized Provision: The diversity hypothesis assumes that public goods are local not just in their benefits but in their origin, and are the product of taxation and spending decisions by the communities that enjoy them. While this pattern may describe some public goods in some countries (and, by construction, laboratory public goods games), it does not describe public goods provision in many real world situations, as some recent work has pointed out (Ejdemyr, Kramon and Robinson, 2015; Singh and vom Hau, 2016). Many public goods are constructed by central or regional governments with little reference to the local community. Highways in the United States, for instance, are usu-

ally constructed and maintained by the state governments rather than the municipalities through which they run. In many other cases, while the goods are provided by a local entity, a large portion of that entity’s funding is provided by higher levels of government, making provision conditional on their decisions. In the US highway example, for instance, much of the money spent by the state government is in fact provided by the federal government. Finally, in some other cases, provision may require independent efforts by both local and supra-local governments. Local governments may build schools but require the central government to staff them with teachers, or regional governments may build power lines but require the central government to build the power plants to make them effective.

While the role of non-local actors is widely acknowledged, with some exceptions authors have tended to treat it as an empirical obstacle to identifying local diversity effects rather than an object of theorizing in itself. To minimize its importance, studies often focus on services in which local initiative is thought to be especially important, or decisions with a specifically local component, such a budget distribution. Some have even argued that the diversity effect holds even for centrally provided goods, since lobbying the non-local government has similar dynamics to cooperative provision (Banerjee and Somanathan, 2007).

The existence of centralized provision, of course, does not mean that local provision does not exist. What it does mean is that finding a situation with “pure” local provision is surprisingly challenging in practice, and that many (if not most) existing observational studies of the diversity effect study processes in which supralocal actors play a role. Alesina, Baqir and Easterly (1999), for instance, studies public goods provision in American cities while acknowledging that the state and federal actors provide considerable funding to these governments, while Banerjee and Somanathan (2007) studies provision in Indian districts while acknowledging that many important distributional decisions are made by state governments.

Powerful Groups: In the classic diversity hypothesis, all groups are modeled as being similar, with only the local sizes of the groups being important. However, different groups may have vastly different social, economic and political positions within the country as a whole. Some groups may gain disproportionate power through their national or regional

numerical size (which may give political advantages in a democratic or quasi-democratic context). In other contexts, the position of the dominant groups may be a result of historical social differences that left groups with legacy of control over land, education and government employment. In the United States, both factors are in play: Whites are not only the majority group, but have historically been overrepresented among the educated and wealthy. In India, by contrast, upper caste members have historically been overrepresented among the educated and wealthy, but in most states are not the plurality group or social category.

The diversity thesis implies such differences in group power should not matter. An area with a minority from the dominant group and a majority from another group (for instance, a city that is 90% black and 10% white) will be expected to have an identical probability of goods provision from an area with a majority from the dominant group and a minority from another group (for instance, a city that is 90% white and 10% black), due to their identical probability of interethnic cooperation. In many countries, this seems a challengeable contention.

Empirically, the problem of differences in group status is compounded by the fact that in many (though certainly not all) cases, the presence of the largest group is associated with fractionalization at the local level. When the powerful group has the largest of share of the population and the number of other groups is small, the correlation between the two measures will be very high by construction, since the largest component of Herfindahl is the square of the largest group's population share. In practice, it might therefore be impossible to distinguish between the negative effects of diversity and the positive effects of the presence of a majority group, due to collinearity.

3 Ethnic Politics and Public Goods: An Alternative Hypothesis

3.1 Bias in Central Government Provision

How might unequally powerful ethnic groups and provision by central institutions be incorporated into a theory of public goods provision? Consider the following simple

scenario. Services, or funds that can be used to purchase services, are provided to local communities by a central government. While the decision-maker is disproportionately likely to be from the powerful group (or is easily influenced by that group), members of that group make up a varying proportion of local communities. The decision-maker, in the presence of a budget constraint, must decide which communities receive goods and funds. There are several reasons to expect that the central government will tend to allocate resources disproportionately to areas inhabited by the powerful group. Two of these are mentioned below:

Intra-ethnic Favoritism: Decision makers (who are disproportionately drawn from more powerful groups) may wish to provide resources to members of their own group. This may be because they have some preference for co-ethnic welfare, face social sanctions when they exclude their own coethnics or tend to attribute positive qualities to members of their own group. All of these have some support from the experimental literature on discrimination and interethnic interaction, which has shown that subjects tend to prefer coethnics in contexts such as voting, even when they are not familiar with the candidate [Adida \(2015\)](#). However, one can obtain a similar predictions from an purely rationalistic theory. In many countries, ethnicity is closely associated with parties and the political system, and politicians use their ethnic identity as a way of getting votes. In Chandra's [\(2007\)](#) influential formulation, in such a system ethnicity serves as a tool for distributing goods to supporters, by serving as a visible signal of political affiliation. This echoes other findings on the association between ethnic politics and clientelism [Kitschelt and Wilkinson \(2007\)](#), and broader theories of the relationship between voter choice and distribution ([Dixit and Londregan, 1996](#)).

The empirical literature on coethnic favoritism in the provision of public services is very extensive. Much of it has focused on the ethnicity of Africa, generally using over-time changes in leader ethnicity as an identification strategy ([Franck and Rainer, 2012](#); [Kramon, 2013](#); [Burgess, Jedwab and Miguel, 2015](#); [Bates, 1983](#); [Lindberg, 2010](#)), though similar studies have focused on India ([Besley et al., 2004](#)) and the world as a whole ([Hodler and Raschky, 2014](#)). This work has shown with some consistency that politicians target resources to members of their own ethnic group, and that voters in turn associate coethnic politicians with superior provision ([Carlson, 2015](#))

Note that ethnic favoritism arguments that ignore intergroup inequality (like most of those above) do not contradict the diversity hypothesis—in fact, being concerned with distribution, they make no predictions about aggregate provision. If the political class is fairly representative of the population, ethnic favoritism will lead to temporal or spatial differences in provision, but in the aggregate each group will get their share, either in districts where their coethnic is elected or in years when their coethnic is elected. In fact, favoritism could provide an alternative causal pathway for the negative effect of diversity, by making homogenous communities more attractive targets (Ejdemyr, Kramon and Robinson, 2015). However, if (as Section Two proposed) decision-makers are *always* more likely to come from wealthy groups, we should expect areas inhabited by these groups to always have better access to public goods.

Other Forms of Ethnic Advantage: However, the advantage enjoyed by dominant groups may extend beyond ethnic favoritism. Members of the most powerful ethnic group may 1) have access to a larger and more powerful social network (since social networks often follow ethnic lines) (Anderson, Francois and Kotwal, 2015). They may also 2) benefit from positive stereotypes about their behavior that were created to justify the status quo power distribution (Jost et al., 2005). Finally, 3) members of particular ethnic groups may have higher levels of political participation and political awareness even after accounting for their higher socioeconomic status (Leighley and Vedlitz, 1999). If any of these factors hold, members of powerful groups may be more likely to gain resources even when a coethnic is not the decision maker, either because they are more likely than members of other groups to be able to influence an influencer of the decision maker, be perceived positively by the decisionmaker, or simply more likely to know how to lobby effectively.

These mechanisms are not mutually exclusive, and they may well act in concert with each other. All of them, however, produce the same basic empirical prediction: When the central government plays a large role in public goods provision, areas with a high proportion of the population from socially and political powerful groups will have higher levels of public goods provision than other areas. Diversity, on the other hand, should have little relationship to goods provision in this situation, since a decision-maker bent on benefiting coethnics will not be sensitive to the internal composition of the non-coethnic

groups.

3.2 Bias by a Non-Democratic Local Governments

Section 3.1 showed how areas with a higher proportion of powerful groups could receive more public goods if the central government plays a large role. While patterns of provision under democracy are the main focus of this paper, it is worth noting that the same predicted results (bias toward areas inhabited by powerful groups) might be observed even if public goods provision is local, so long as the local government is not democratic. Recall that the theory assumed that the “central” decision-maker was politically distinct from the local community and susceptible to influence by the most powerful groups in the region. This may occur at the local level as well, either when a local government is not elected or is able to deter members of non-prominent groups from voting. In the Jim Crow South, for instance, repression of black voter participation enabled whites to control local governments even in areas where they were a minority.

Such non-elected local elites may pursue a variety of policies for different types of goods, but there are three reasons to expect that they will be especially hesitant to expand the provision of truly non-excludable goods in areas where they are in the minority. Firstly, they may be unwilling to expend effort or tax money on projects that will benefit mostly non-coethnics, following existing experimental findings in non-cooperative games (Fehr, Hoff and Kshetramade, 2008). Secondly, in areas where the elite have few coethnics they may be less exposed to the type of informal pressures for public goods provision described by Tsai (2007). Thirdly, elites may prefer to have lower levels of public goods provision if they think higher levels will provide the excluded group with the resources to challenge their control in a subsequent period (Anderson, Francois and Kotwal, 2015).

3.3 Discussion

In cases where the dominant group is also the largest group, the ethnic discrimination hypothesis described in Section 3.1 produces predictions very similar to the diversity model. The areas populated by the most powerful group, which will also tend to be

the less diverse areas, will have higher levels of public goods provision. The theoretical implications, however, are quite different, since one emphasizes central discrimination and the other local cooperation. To understand the differences, consider the case of the city of Flint, Michigan, which combines a non-white population and fractionalization index that are high relative to other cities in Michigan with a set of public services (most notoriously, its water system) that are of much lower quality. From the diversity perspective, this pattern would be attributed to the inability of actors within Flint to cooperate to provide high quality services. From the discrimination perspective, this pattern would be attributed to a pervasive set of biases within the state and federal government that lead to lower levels of resources being provided to poor cities in general, and poor non-white cities in particular.¹

This does not necessarily mean that all goods will be provided in a discriminatory fashion. For many goods in many countries, there are strong technocratic criteria that guide distribution. If these programs operate as designed, it may even lead central transfers to flow disproportionately to disadvantaged communities. For other goods, however, particularly those provided through less transparent or more politicized procedures, the rules may give considerable discretion for central decision-makers to favor the powerful.

These explanations are also not mutually exclusive: The discrimination logic might apply to cases where goods are provided by the central government and there are strong power asymmetries within groups, while the diversity logic might apply in cases where groups are relatively equal and local governments relatively autonomous. However, given the prevalence of ethnic inequalities and overcentralized systems of local government in both the world in general and developing countries in particular, the discrimination perspective is worthy of sustained empirical investigation.

¹While this theory is similar to existing work on the importance of intergroup inequality (Baldwin and Huber, 2010), it propounds a distinct mechanism, claiming that intergroup inequality allows specific groups to win a large share of public benefits, rather than that intergroup inequality makes cooperation more difficult. Baldwin and Huber (2010) also examine national rather than local outcomes.

4 The Indian Case: Data

4.1 Caste in India

To separate these two mechanisms, it is necessary to examine a case where the proportion of individuals from the highest status groups is uncorrelated with the fractionalization. For this reason, the main analysis will focus on public goods provision in rural Northern India. All regions of peninsular India have some degree of exposure to the caste system, a ranked system of identities which, at least in theory, assigns relative status to India's hundreds of endogamous caste groups, or *jatis*. In popular discussion, *jatis* are often discussed in terms of the categories to which they are assigned for the purpose of affirmative action: "upper" or "twice born" castes; "intermediate" or "peasant" castes; "other backward classes" and "untouchables" or "scheduled castes." The "scheduled tribes" of the hills and the Muslim minority are somewhat outside of this system, but are usually considered to be poor and politically marginalized. These levels of social status are associated with, though not identical to, levels of wealth, education and traditional political power. Certain groups that combine relatively high status, large size, and control over land are referred to as "dominant castes", and exercise disproportionate political power within specific regions (Frankel and Rao, 1989; Srinivas, 1987; Lee, 2016).

Caste differs in several important respects from ethnic identities in other parts of the world, especially in its embrace to ranking. However, none of these differences lead to violations of the basic assumptions about ethnic divisions that underlie the diversity hypothesis: That goods preferences be heterogeneous across groups, that social sanctioning be more effective within groups, and that identities be commonly known. In fact, experimental evidence in India shows that there are great difficulties in cooperation between castes and low levels of other-regarding preferences across castes (Fehr, Hoff and Kshetramade, 2008). Similarly Banerjee and Somanathan (2007) replicate the diversity finding across Indian districts, leading many to cite India as support for the diversity hypothesis.²

²For a more detailed defense of the use of caste to test hypotheses in the ethnic politics literature, see Chandra (2007, 16-18).

There is considerable evidence that caste plays a role in public service distribution decisions in India. Not only does the caste identity of officials appear to effect who received services (Besley et al., 2004; Pande, 2003), but upper caste individuals in particular appear to have strong preferences against any type of redistribution to the lower castes (Fehr, Hoff and Kshetramade, 2008). A perception of discrimination by the upper castes has also been at the center of the movements for lower caste empowerment that have flourished since the early 20th century (Jaffrelot, 2003).

Importantly, the presence of upper and intermediate castes is uncorrelated with the overall fragmentation of the caste system. Not only are Hindu upper and intermediate castes a relatively small proportion of the population (27%, according to the 2005 National Family Health Survey), but the category is internally divided into a variety of jatis, such as Brahmins, Rajputs, Jats and Marathas. Overall, no Indian jati comprises more than 38% of the population of a major state (Lee, 2016). It is thus very possible for an area to be either very diverse with a high upper caste percentage (if several high status jatis are present) or homogenous with a low upper caste proportion (if one of the several relatively large low-status jatis is present).

4.2 The Data

The data used in this paper are drawn from the UP-Bihar Survey of Living Conditions, conducted by the World Bank in northern and central Bihar and eastern Uttar Pradesh in 1997 and 1998. These two states are among the poorest areas of India, with reputations for low public service quality. The data thus show considerable variation in provision of even very basic public goods. These areas are also well known for the high salience of caste in politics, making both the diversity and discrimination mechanisms potentially relevant.

The UPBSLS covered 120 villages, evenly divided across the two states, and selected using a stratified random design built upon an earlier qualitative survey. For each village, the enumerators collected a large amount of information on overall access to public services and social conditions, while also conducting a detailed individual survey of 30 (in some villages, 15) randomly chosen households. The USBLS was used here

despite its age because the village-level data includes the population breakdown of castes in the village as a whole. This makes the survey a valuable rarity among village-level studies in India, given the government’s reluctance to collect data on caste.

The village-level caste enumeration departed in two key ways from a population census, since it only included the seven most populous castes in the village (for reasons of space on the form) and enumerated households rather than individuals. To deal with this problem, the main models use as a population measure the proportion of households of each caste within the total number of households from the seven largest castes. A variety of alternative approaches to this issue are discussed in an online appendix, and shown to have no impact on the results.

4.3 Unequal Caste Power

In the study area, both the anthropological literature and the available data show that four castes have traditionally had disproportionate levels of social and political influence: The Brahmins, Bhumihars, Rajputs and Kayasths, derisively referred to by their opponents by the partial acronym *bhurabal*. All four groups are of high ritual status (“twice born”), and are of much higher ascribed status than any other local group. These four groups are also notable for their economic and political influence. Collectively, they comprise 15% of the population of Bihar, but own 48% of the land and provide 44% of the state’s college graduates (Lee, 2016). They have also traditionally been powerful within the political system and the bureaucracy. In 2005, they provided 26% of Bihar’s state legislators (Jaffrelot and Kumar, 2012). In 2012, the four dominant castes provided 77% of Bihar senior civil servants, despite strict quotas (Witsoe, 2013, 85). Qualitative accounts of the politics and society of Bihar and UP also emphasize the enduring political and social importance of these groups, founded on their high levels of wealth and education (Witsoe, 2013; Jha, 1972; Mishra and Pandey, 1996; Chakravarti et al., 2001).

Is this regional dominance likely to be helpful to the relatively poor and rural members of the upper castes in our sample? Evidence from the Indian Human Development Survey (2005) shows that in these regions even ordinary upper caste villagers are much more likely to have useful contacts in the state apparatus than their lower caste neigh-

bors. Table A.9 shows that relative to other caste members, upper caste respondents are more than twice as likely to know a government employee and three times as likely to know a government officer. Some, though not all, of this difference seems to be driven by caste-specific social networks, since the contacts of upper caste members are 25 percentage points more likely to come from their own caste as the contacts of others.

Upper caste power in these regions has been so marked that it has led to an compensating movement for lower caste self-assertion. Riding this wave, lower caste politicians such as Mayawati, Laloo Prasad Yadav and Mulayam Singh held the highest political offices in both states during periods of the 1990s. However, there are a variety of reasons to doubt that this change, still somewhat novel and insecure in 1997, would have eliminated the relationship between upper caste interests and distributional patterns. Firstly, the majority of public goods at the time of the survey would have been constructed before 1990, under the influence of upper caste politicians. Secondly, upper caste political involvement remains very high, both within the civil service and their role in the newly important lower-caste-led political parties (Jaffrelot and Kumar, 2012). Finally, most accounts of the these lower caste politicians have emphasized the degree to which they neglected public goods provision in favor of rent-seeking and patronage distribution, to a degree even more notable than their upper caste predecessors (Witsoe, 2013; Thakur, 2000).

However, despite their political power, the small size of the twice born groups mean that their presence in villages is unassociated with diversity: Overall, the Herfindahl index of caste diversity of villages in the sample has a correlation with percent upper caste of $\rho = -.01$.³

4.4 The Dependent Variable

The UPBSLS has information on the actual incidence of a wide variety of public amenities available in each village. Four are the focus of the empirical analysis here: One continuous measure (the number of hours of electricity the village receives in a “good

³Following standard practice, the Herfindahl index of caste fractionalization for each village is calculated as $Fractionalization = 1 - \sum_{i=1}^n p_i^2$, where p_i is the share of households of the top seven jatis of jati i . Muslim social groups are treated as castes.

month”) and three binary measures (the presence of a public primary school, public primary health clinic, and landline telephone access. All are, at least potentially, broadly beneficial to residents of a village that has them and similar (if not identical) to the types of local public goods discussed in the broader literature. All are also provided by the state in India, with no possibility (by definition, in most cases) of private substitution.

These particular goods were selected because of their prominence in existing accounts of distributive politics in India. Schools, clinics, and telephones were all used as outcomes in [Banerjee and Somanathan \(2007\)](#), the most influential Indian formulation of the diversity hypothesis. Hours of electricity was chosen because [Min \(2015\)](#) provides one of the best existing accounts of the microprocesses by which state officials in the study area (in this case power engineers) distort public goods provision to favor areas with more political and economic clout. Min notes that “politicians routinely interfere in the operation of the state electricity board, from patronage transfers of employees, interventions in the selection of villages for electrification projects, and the assertion of influence on when, where, and how, power cuts are timed and distributed.” ([Min, 2015](#), 130).

4.5 The Distributional Process

The political and economic order that supplies these goods represents a mix of state and local institutions and pressures. Power lines, schools, and health clinics may be constructed by either the state government or the village (panchayat) government. The state government provides staff and funding for clinics and schools, and service, maintenance, current for phones and electricity, and a large share of the funding of the local panchayat. The panchayat is responsible for the maintenance of many of these same facilities. The local community, in addition, must lobby hard to receive state resources, and much of the activity of the panchayat government is taken up in this type of maneuvering. [Banerjee and Somanathan \(2007\)](#) suggest that cooperative lobbying is an important conduit for the diversity hypothesis.

This complex distributional arena might encourage proponents of both hypotheses discussed above. While [Banerjee and Somanathan \(2007\)](#) treated the provision process

as being fundamentally local (and thus a good place to test the diversity hypothesis), others might argue that the power of the state government (and the power of the upper castes in that government) would make discrimination the obvious mechanism. While the Indian pattern differs considerably from the ideal types of either central or local provision, it is typical of the type of messy, multi-level distributional process common in many parts of the world. In Section Six we will discuss several non-Indian examples of such processes.

5 The Indian Case: Results

5.1 The Diversity Hypothesis

The determinates of public goods provision in UP and Bihar are examined in Table 1, which shows the results of a series of regression models, using linear regressions for the number of hours of electricity and logistic regressions for the other three goods. All models control for the population of the village, since population is an obvious “technocratic” predictor of whether a village should be targeted for service delivery by the state. Population is also mentioned by Gerring et al. (2015) as a key factor in explaining variation in the effect of ethnicity on goods provision. Panel B of Table A.7 shows that the results are robust to not including this variable. In the appendix, a variety of alternative models are reported, including one using district fixed effects.

Panel A of Table 1 examines the association between caste fractionalization in a village and a village’s level of public services. There is no evidence of a strong negative relationship between fractionalization and public services. In fact, if anything, the results go the other way, with the effect of fractionalization being statistically significant and positive for both telephones and health centers. This provides some evidence that, at least within India, local diversity is not a negative factor for obtaining services.

One longstanding critique of the diversity hypothesis is that it does not account for unobserved differences in geography and economics between diverse and homogenous places. Panel B of Table 1 includes a vector of controls designed to capture such differences. These include dummy variables for whether the village was in Bihar or had road

Table 1: The Diversity Hypothesis: Public Goods in India

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Fractionalization</i>				
Village Caste Fract.	1.451 (2.397)	3.578* (1.903)	-0.841 (1.258)	11.97** (5.334)
Village Pop.	0.00316 (0.00289)	0.00157 (0.00131)	0.0108*** (0.00311)	0.00309* (0.00159)
Constant	2.417 (1.628)	-4.038*** (1.345)	-0.300 (0.871)	-11.66*** (4.081)
Observations	120	119	119	119
<i>Panel B: Controls</i>				
Village Caste Fract.	0.594 (2.411)	4.542* (2.330)	-0.555 (1.418)	14.24** (6.265)
Village Pop.	0.00507* (0.00296)	0.000576 (0.00148)	0.0116*** (0.00335)	0.00320 (0.00199)
Bihar	-1.476 (1.138)	0.758 (0.603)	2.024** (0.786)	1.217 (0.949)
% House. With Off-Farm Jobs	0.0194 (0.0180)	0.000621 (0.0106)	0.00704 (0.00967)	0.0191 (0.0170)
Village Road Access	2.692* (1.440)	1.350 (1.114)	-0.676 (0.933)	-1.034 (1.038)
Village Prop. Landless	-0.00690 (0.0270)	0.00794 (0.0141)	-0.0218 (0.0179)	-0.00333 (0.0231)
Irrigation % Scoring	0.614 (0.392)	0.126 (0.229)	0.388* (0.225)	0.692* (0.409)
Avg. Price No Irrigated Land	9.82e-06 (1.26e-05)	2.39e-06 (6.49e-06)	-2.49e-06 (7.28e-06)	2.83e-06 (9.02e-06)
Constant	-2.635 (2.702)	-7.037*** (2.309)	-1.929 (1.709)	-16.84*** (5.425)
Observations	116	115	115	115

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

access (paved or unpaved), the proportion of the land that was irrigated, the average price of unirrigated land, and the proportion of households that were landless and had off the farm jobs. The inclusion of these controls does not alter the basic results: Caste fractionalization has no negative effect on public goods provision.

5.2 The Discrimination Hypothesis

Table 2 examines the effect of the proportion of upper caste households on service provision. Panel A tests the simple association, conditional on village population. The proportion of upper castes has a strong, statistically significant positive association with public goods provision. The increase from an entirely lower caste village to an entirely upper caste one would increase the predicted number of hours of electricity a village receives in a week by four hours (a whole standard deviation) while the same change for a village with the median population would increase its predicted probability of having a public primary school from 77% to 97%.

Panel B adds measures of caste fractionalization to the model. While caste fractionalization still does not have a negative effect on goods provision, upper caste presence retains a statistically significant and positive effect. At least within India, the presence of specific groups appears more closely associated with public services than overall diversity.

Panels C and D test the robustness of the association between upper caste presence and public goods. Panel C includes the standard vector of economic controls, which has little effect on the upper caste coefficient. Panel D controls for the presence of two especially disadvantaged types of social groups, Muslims and Scheduled Castes and Tribes, with castes classified as Other Backward Classes (OBCs) as the excluded category. The presence of neither of these groups has an effect on public goods provision statistically different than the OBCs, confirming that the results are capturing an upper caste advantage rather than the disadvantages of some other particular category.

Table 2: The discrimination hypothesis: Public Goods in India

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Upper Caste Population</i>				
Village Pop.	0.00387 (0.00278)	0.00243* (0.00127)	0.0120*** (0.00328)	0.00419*** (0.00152)
Village Prop. Upper Castes	4.345* (2.274)	2.307** (1.014)	3.447** (1.670)	2.146* (1.290)
Constant	2.577*** (0.917)	-2.200*** (0.476)	-1.438** (0.659)	-3.627*** (0.675)
Observations	120	119	119	119
<i>Panel B: Upper Caste Population and Fractionalization</i>				
Village Pop.	0.00348 (0.00286)	0.00164 (0.00135)	0.0125*** (0.00338)	0.00300* (0.00161)
Village Prop. Upper Castes	4.346* (2.281)	2.916** (1.154)	3.685** (1.801)	3.580** (1.799)
Village Caste Fract.	1.456 (2.371)	4.393** (2.113)	-1.025 (1.316)	13.09** (5.426)
Constant	1.736 (1.649)	-5.101*** (1.565)	-0.908 (0.944)	-13.07*** (4.239)
Observations	120	119	119	119
<i>Panel C: Controls</i>				
Village Pop.	0.00586** (0.00284)	0.00174 (0.00143)	0.0130*** (0.00363)	0.00505** (0.00197)
Village Prop. Upper Castes	5.033** (2.263)	2.061* (1.110)	3.265* (1.739)	2.592* (1.531)
Controls	Yes	Yes	Yes	Yes
Constant	-3.313 (2.308)	-4.511*** (1.558)	-2.749* (1.617)	-7.162*** (2.083)
Observations	116	115	115	115
<i>Panel D: Other Caste Categories</i>				
Village Pop.	0.00397 (0.00284)	0.00247* (0.00129)	0.0126*** (0.00345)	0.00466*** (0.00159)
Village Prop. Upper Castes	4.230* (2.467)	2.230** (1.117)	2.996* (1.702)	2.957* (1.587)
Village Prop. Scheduled Castes/Tribes	0.189 (2.454)	0.0119 (1.269)	0.00353 (1.219)	2.564 (1.746)
Village Prop. Muslim	-0.465 (2.299)	-0.286 (1.239)	-1.499 (1.210)	0.776 (1.742)
Constant	2.571* (1.358)	-2.170*** (0.697)	-1.310 (0.861)	-4.741*** (1.137)
Observations	120	119	119	119

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable. The controls in Panel C are the same as those in Panel B of Table 1.

5.3 Identity Salience and Social Inequality

Many formulations of the diversity hypothesis have rejected the idea that identity fractionalization is always associated with bad outcomes, arguing instead that this effect is conditional of the effects of ethnic identity being salient. Alternatively, ethnicity might become a focus of conflict if there are high levels of economic inequality between groups, making differences in behavior or redistributive preferences especially large. In the Indian context, this would mean that the relevant variable for the diversity hypothesis is not caste fractionalization but social distance between castes, intercaste inequality, or the interaction of the two with fractionalization. Section A.2 tests these hypotheses, and finds that even after accounting for the endogenous salience of social difference, the presence of upper castes has a strong influence on the presence of public goods.

5.4 Alternative Hypotheses: Wealth, Social Capital and Patronage

Wealth: Perhaps the most obvious alternative explanations for the positive relationship between public goods is economic differences between villages. Wealthier villages might both have more upper caste residents and be better able to afford high quality public services without central assistance. While the geographical controls in Tables 2 and 1 address some of these concerns about unobserved differences between villages, they do not measure differences in wealth and income directly.

Measuring wealth in rural India is difficult, due to the prevalence of non-cash income and transfers. To deal with this problem it is customary to estimate income using expenditures or household stocks of durable goods. Table A.10 includes three such measures, each calculated from the household sample within each village: Average aggregate household expenditure in the previous 30 days (with non-cash expenditures estimated), average values of the scores of the first dimension of a factor analysis of the possession of 15 household durable goods, and the proportion of households living in “pucca” [well-constructed] houses, as judged by the survey taker. While wealth is generally positively associated with the availability of public goods, the inclusion of the variable does not effect of the positive, statistically significant effect of upper caste presence, except for

electric power in some models.⁴

Intra-Elite Cooperation and Social Capital: Section Three suggested that the association between the presence of powerful groups and public goods is driven by the superior ability of these groups to obtain favors from higher levels of government. There is, however, an alternative explanation for these findings—that members of politically powerful groups are better at providing public goods for themselves, due to higher levels of social capital and social skills than other groups. While this pathway would be distinct from the diversity hypothesis, it would emphasize the role of local initiative and superior ability to resist free-riding, rather than non-local policy. This mechanism would be even more likely if we allow the possibility (considered in the appendix) that local public good can be made excludable.

There is reason to doubt that higher within-group cooperation is driving the results. Many anthropological accounts of village India stress the importance of factional divisions within upper caste groups, divisions that are often more politically relevant than caste divisions (Lewis, Barnouw et al., 1958; Miller, 1965).

Testing the social capital measure directly is hampered by the UPBSLS’s not asking questions that directly pertain to intracaste (or intercaste) cooperation. To gain some perspective on this issue, Table A.9 presents levels of a variety of social capital measures taken from the 2005 Indian Human Development Survey, with the sample truncated to correspond to the UPBSLS’s sampling area. There appears strong evidence that conventional measures of social capital are not higher among the upper castes. Membership in caste associations is actually higher among the lower castes, as are (by a smaller margin) membership in other private associations and generalized preferences for local collective action. Interestingly, the only type of associational membership more common among the upper castes is the only one sponsored by the state, cooperative societies, which in this region serve as a way for channeling cheap loans to the politically well-connected (Witsoe, 2013).

Patronage: A third alternative hypothesis is that the poor public goods are a result of widespread patron-client networks, and the desire of patrons to maintain these networks.

⁴Some of this effect might be driven by the endogeneity of public goods and wealth. Certain of the household durable goods, for instance, require electric power to operate.

Patrons (who, in the South Asian context, are likely to be upper caste) use their control over land and the risk-avoidant behavior of the poor to keep their clients in a dependent position. Public services such as education, which expands the outside options of clients, are thus against their interests, and powerful patrons will prevent their implementation, leading to suboptimal outcomes (Anderson, Francois and Kotwal, 2015; Shami, 2012).

While this is a plausible and well-attested mechanism, there are reasons to think that it cannot explain the strong association between upper caste presence and public goods. Firstly, we should expect any patronage effect to be concentrated among goods specifically targeting the poor, and affecting wages: In fact, Anderson, Francois and Kotwal (2015) find no association between their clientelism measure and the broadly beneficial goods discussed here. Secondly, if we assume that villages with an upper caste presence are more likely to exhibit clientelistic tendencies, we should expect these villages should receive less public goods rather than more.

Table A.11 examines the association between upper caste population and public goods while conditioning by two direct measures of upper caste patronage taken from the household part of the UPBSLS. The first is the proportion of loans to lower caste households given by upper caste individuals other than shopkeepers, and the second is the proportion of lower caste households where the primary source of income was casual agricultural labor, usually identified as being the groups most vulnerable to exploitation by patrons. Neither measure has a noticeable effect on public goods provision, while the effect of the village's upper caste population remains positive.

6 Non-Indian Cases

Given the prevalence of results showing a strong connection between diversity and public goods, the results in Section Five showing no connection might be thought to stem from some idiosyncratic feature of the Indian case, such as the relatively small roll of local initiative in public goods and the hierarchical nature of the caste system. To examine the wider generalizability of these findings, this section examines this hypothesis, using data from two influential studies of the diversity hypothesis: Alesina, Baqir and Easterly (1999) and Miguel (2004). The results support the claim made in Section Two: The

multicollinearity makes it difficult to separate the diversity and power hypotheses in most cases.

6.1 Urban America: Replicating Alesina, Baqir and Easterly (1999)

Alesina, Baqir and Easterly (1999) [ABE], in one of the first studies of the ethnicity on public goods spending, found a negative relationship between racial fractionalization and spending on several categories of public services in American cities and counties. They attribute this problem to heterogeneous preferences across racial groups, which lead many voters to have preferences over goods very different from the median voter, making them unwilling to tax themselves to fund the median voter’s preferred projects.

In many ways, urban America should be a more favorable case for the diversity hypothesis than rural India, given the relative autonomy of American local institutions. However, even in the United States, a large proportion of local budgets come from federal and state funds, funds that often come with constraints on how the money may be spent. While ABE attempt to account for this spending in their models, it is unclear if they can properly separate formulaic transfers to poorer communities from more discretionary favoritism. It is particularly notable that ABE’s strongest results come from a sector (roads) widely reputed to involve “pork” transfers, and that they do not find a direct effect of diversity on local taxes.

Another feature of race in America is that, relative to caste in India, the number of groups is smaller and the population of the most politically powerful group, white Americans, is much larger. This means that the correlation between diversity and the population of white Americans is very high: In ABE’s city sample, $\rho=.86$. The consequences of this close association are apparent in Table 3. Panel A reproduces ABE’s core results for their sample of US cities: Expenditure on roads and sewage is negatively correlated with diversity, while expenditure on police rises. Panel B shows that the same associations hold for public goods and the proportion of white city residents. However, when both variables are included together (Panel C) the results become less clear. The estimated effect of fractionalization on relative road spending remains stati-

cally significant and negative, while the effect of white population remains statistically significant and positive. However, for relative sewage and per capita roads expenditure both estimated coefficients are diminished and statistically insignificant.

Table 3: Replicating Alesia, Baqir and Easterly (1999): Ethnicity and Public Goods in US Cities

VARIABLES	(1) Prop. Police	(2) Prop. Roads	(3) Prop. Sewage	(4) Road Exp. PC
<i>Panel A: ELF</i>				
Fractionalization	0.0566*** (0.0118)	-0.0962*** (0.0114)	-0.0488*** (0.0160)	-35.50*** (9.027)
Constant	0.139*** (0.00403)	0.137*** (0.00388)	0.143*** (0.00546)	91.00*** (3.079)
Observations	1,020	1,020	1,020	1,020
R-squared	0.022	0.066	0.009	0.015
<i>Panel B: Majority Ethnic Group</i>				
Prop. White	-0.0478*** (0.0123)	0.101*** (0.0118)	0.0432*** (0.0166)	39.36*** (9.341)
Constant	0.193*** (0.00999)	0.0285*** (0.00958)	0.0941*** (0.0135)	49.34*** (7.596)
Observations	1,020	1,020	1,020	1,020
R-squared	0.015	0.067	0.007	0.017
<i>Panel C: ELF and Majority Ethnic Group</i>				
ELF	0.0777*** (0.0268)	-0.0468* (0.0257)	-0.0587 (0.0363)	-7.227 (20.43)
Prop. White	0.0244 (0.0277)	0.0570** (0.0267)	-0.0114 (0.0376)	32.64 (21.16)
Constant	0.113*** (0.0293)	0.0768*** (0.0282)	0.155*** (0.0398)	56.79** (22.39)
Observations	1,020	1,020	1,020	1,020
R-squared	0.023	0.070	0.009	0.017

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The models report the results of a series of linear regression models. Compare to [Alesina, Baqir and Easterly \(1999\)](#), Table 4.

These results do *not* show that the effect of the presence of a socially powerful group (whites) is more important than fractionalization. Rather, they show that the two measures are so highly collinear that it is impossible to separate their effects. This emphasizes the importance of the research design used in this paper, where diversity can be separated from the presence of a single group. Combined with what we know about the political economy of local finance in the United States, it suggests that differences

between white and non-white cities are at least *potentially* significant, even in cases cited as evidence for the diversity hypothesis.

6.2 Rural Kenya: Replicating Miguel (2004)

Miguel (2004) found that village-level ethnic diversity had strong effects on the quality of public services in rural Kenya (where ethnicity is highly politically salient) but had no such effect in otherwise comparable parts of Tanzania (where ethnicity is less salient). Since Miguel only finds a diversity effect in Kenya, this section only reproduces the Kenyan results. Like ABE, Miguel notes that the provision of public services in Kenya is a complicated, multi-level affair, involving both local communities and the central government: In fact, Miguel states that 90% of funding for schools (his central outcome) comes from the central government, though local funding might be important on the margins, since “local inputs and teachers are complements in educational production.” However, given the strong role of non-local actors in the production of public services, non-cooperative mechanisms certainly seem worth examining in the Kenyan context.

As in the United States, the areas of Western Kenya in this study have a majority group, the Luhya. While the Luhya are neither the largest or wealthiest group in the country as a whole, they are considered the “dominant” group locally, and have been associated with the region for longer than the other major groups Miguel (2004). Overall, Luhya presence and ethnic fractionalization are highly correlated, at $\rho=.64$.

Panel A of Table 4 replicates Miguel’s key findings: Ethnic diversity is negatively associated with the provision of a variety of public services, a relationship that is statistically significant for school funding, school desks, and well maintenance. However, Panel B shows that very similar results can be obtained by substituting the proportion of Luhya in the village population, which is positively associated with the same goods. When the two variables are included in the same model (Panel C) both sets of coefficients diminish sharply, while standard errors become much larger. Both these effects, however, are less pronounced for the Luhya coefficient estimates, which remain statistically significant in the combined model.

As with the ABE results, collinearity makes it difficult to separate the effect of

Table 4: Replicating Miguel (2004): Ethnicity and Public Goods in Rural Kenya

VARIABLES	(1) School Funds	(2) Desks Per Pupil	(3) Latrines Per Pupil	(4) Classrooms Per Pupil	(5) Prop. Working Wells
<i>Panel A: ELF</i>					
ELF	-3.528** (1.739)	-0.291*** (0.0864)	-0.0152 (0.0106)	-0.00981 (0.0113)	-0.325*** (0.110)
Constant	4.278*** (0.473)	0.276*** (0.0235)	0.0196*** (0.00289)	0.0329*** (0.00307)	0.632*** (0.0300)
Observations	84	84	84	84	84
R-squared	0.048	0.121	0.024	0.009	0.096
<i>Panel B: Majority Ethnic Group</i>					
Prop. Luhya	1.753*** (0.659)	0.153*** (0.0313)	0.00136 (0.00415)	0.00475 (0.00433)	0.133*** (0.0422)
Constant	2.314*** (0.487)	0.109*** (0.0231)	0.0151*** (0.00306)	0.0275*** (0.00320)	0.470*** (0.0312)
Observations	84	84	84	84	84
R-squared	0.079	0.226	0.001	0.014	0.108
<i>Panel C: ELF and Majority Ethnic Group</i>					
ELF	-1.215 (2.160)	-0.0792 (0.102)	-0.0206 (0.0134)	-0.00368 (0.0142)	-0.184 (0.137)
Prop. Luhya	1.469* (0.833)	0.134*** (0.0395)	-0.00347 (0.00518)	0.00389 (0.00548)	0.0899* (0.0529)
Constant	2.784*** (0.968)	0.139*** (0.0458)	0.0231*** (0.00601)	0.0289*** (0.00637)	0.541*** (0.0614)
Observations	84	84	84	84	84
R-squared	0.083	0.231	0.030	0.015	0.127

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The models report the results of a series of linear regression models. Compare to Miguel (2004), Table 2.

diversity from the presence of the local majority group in the Kenyan case. They do, however, suggest that the discrimination hypothesis is just as consistent with the data as the diversity hypothesis, and in fact seems to fit the variation somewhat better.

7 Conclusion

The results presented in Sections Five and Six challenge the idea that the link between ethnic identity lies in the failure of ethnic groups to cooperate. Instead, they suggest that certain socially and politically powerful ethnic groups are better able than others to obtain services for their communities than others, whether because of their coethnic links to higher level political actors or the generally stronger social position of group members . While these results suggest a different mechanism from that outlined in the diversity literature, they are consistent with the data from much of this literature, due to the collinearity between the presence of powerful groups and standard diversity measures.

This is not to say that cooperation, and failures of cooperation caused by ethnic divisions, do not occur. In some situations, such as the provision of goods in highly decentralized countries with limited ethnic hierarchy, the cooperation mechanism may very accurately capture the process by which public services are supplied. However, for many services in many countries, provision involves actors outside local communities, and some groups are more influential than others in the polity as a whole. In these situations, bias toward powerful groups is be an important factor in distribution, consistent with the large existing literature on ethnic favoritism.

Since they focus on distributional decisions within countries, these results cannot speak to the relationship between diversity and public goods provision at the national level. However, the results complement recent work that shows that country-level inter-ethnic inequalities are associated with failures of public goods provision ([Baldwin and Huber, 2010](#)) by showing how national-level conflicts can influence local level-variation.

These results have implications for policy. Many development interventions, motivated by the diversity literature, have sought to improve public services by improving within-village cooperation ([Fearon, Humphreys and Weinstein, 2009](#)). These findings

suggest, by contrast, that a focus on the policies of regional or national political actors is often appropriate. This is not to suggest the behaviors of these actors is easy: Bias toward the powerful and wealthy is deeply seated in most political systems. However, a good place to start is by accepting that the problems of diverse communities are, at least in some cases, beyond their control.

Bibliography

- Adida, Claire L. 2015. "Do African voters favor coethnics? Evidence from a survey experiment in Benin." *Journal of Experimental Political Science* 2(01):1–11.
- Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics* 114(4):1243–1284.
- Algan, Yann, Camille Hémet and David Laitin. 2011. "Diversity and public goods in a republican setting." *Journal of Political Economy* .
- Anderson, Siwan, Patrick Francois and Ashok Kotwal. 2015. "Clientelism in Indian villages." *The American Economic Review* 105(6):1780–1816.
- Baldwin, Kate and John D Huber. 2010. "Economic versus cultural differences: Forms of ethnic diversity and public goods provision." *American Political Science Review* 104(04):644–662.
- Banerjee, Abhijit, Lakshmi Iyer and Rohini Somanathan. 2005. "History, social divisions, and public goods in rural India." *Journal of the European Economic Association* 3(2-3):639–647.
- Banerjee, Abhijit and Rohini Somanathan. 2007. "The political economy of public goods: Some evidence from India." *Journal of Development Economics* 82(2):287–314.
- Bates, Robert H. 1983. Modernization, ethnic competition, and the rationality of politics in contemporary Africa. In *State versus ethnic claims: African policy dilemmas*. Vol. 152 Boulder, CO: Westview Press p. 171.

- Besley, Timothy, Rohini Pande, Lupin Rahman and Vijayendra Rao. 2004. "The politics of public good provision: Evidence from Indian local governments." *Journal of the European Economic Association* 2(2-3):416–426.
- Burgess, Robin, Remi Jedwab and Edward Miguel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *The American economic review* 105(6):1817–1851.
- Carlson, Elizabeth. 2015. "Ethnic voting and accountability in Africa: A choice experiment in Uganda." *World Politics* 67(02):353–385.
- Chakravarti, Anand et al. 2001. *Social power and everyday class relations*. Sage Publications.
- Chandra, Kanchan. 2007. *Why ethnic parties succeed: Patronage and ethnic head counts in India*. Cambridge University Press.
- Dixit, Avinash and John Londregan. 1996. "The determinants of success of special interests in redistributive politics." *the Journal of Politics* 58(04):1132–1155.
- Easterly, William and Ross Levine. 1997. "Africa's growth tragedy: policies and ethnic divisions." *The Quarterly Journal of Economics* pp. 1203–1250.
- Ejdemyr, Simon, Eric Kramon and Amanda Lea Robinson. 2015. "Segregation, Ethnic Favoritism, and the Strategic Targeting of Distributive Goods¹." APSA Annual Meeting Paper <http://stanford.edu/~ejdemyr/docs/EKR-20151103.pdf>.
- Fearon, James D, Macartan Humphreys and Jeremy M Weinstein. 2009. "Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia." *The American Economic Review* 99(2):287–291.
- Fehr, Ernst, Karla Hoff and Mayuresh Kshetramade. 2008. "Spite and development." *The American Economic Review* pp. 494–499.
- Franck, Raphael and Ilia Rainer. 2012. "Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa." *American Political Science Review* 106(02):294–325.

- Frankel, Francine R and Madhugiri Shamarao Ananthapadmanabha Rao. 1989. *Dominance and State Power in Modern India: decline of a social order*. Vol. 1 Oxford University Press Oxford.
- Gerring, John, Strom C Thacker, Yuan Lu and Wei Huang. 2015. "Does diversity impair human development? A multi-level test of the diversity debit hypothesis." *World Development* 66:166–188.
- Gisselquist, Rachel M, Stefan Leiderer and Miguel Niño-Zarazúa. 2016. "Ethnic Heterogeneity and Public Goods Provision in Zambia: Evidence of a Subnational "Diversity Dividend"." *World Development* 78:308–323.
- Habyarimana, James, Macartan Humphreys, Daniel N Posner and Jeremy M Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review* 101(04):709–725.
- Hodler, Roland and Paul A Raschky. 2014. "Regional favoritism." *The Quarterly Journal of Economics* p. qju004.
- Jaffrelot, Christophe. 2003. *India's silent revolution*. Orient Blackswan.
- Jaffrelot, Christophe and Sanjay Kumar. 2012. *Rise of the Plebeians?: The Changing Face of the Indian Legislative Assemblies*. Routledge.
- Jha, Shashi Shekhar. 1972. *Political elite in Bihar*. Bombay: Vora.
- Jost, John T, Yifat Kivetz, Monica Rubini, Grazia Guermandi and Cristina Mosso. 2005. "System-justifying functions of complementary regional and ethnic stereotypes: Cross-national evidence." *Social Justice Research* 18(3):305–333.
- Kitschelt, Herbert and Steven I Wilkinson. 2007. "Citizen-politician linkages: an introduction." *Patrons, clients, and policies: Patterns of democratic accountability and political competition* pp. 1–49.
- Kramon, Eric. 2013. *Vote Buying and Accountability in Democratic Africa* PhD thesis Stanford University.

- Lee, Alexander. 2016. "Development in Multiple Dimensions: Social Power and Regional Policy in India." Unpublished Manuscript https://www.rochester.edu/college/faculty/alexander_lee/wp-content/uploads/2014/09/chapter1revised.pdf.
- Leighley, Jan E and Arnold Vedlitz. 1999. "Race, ethnicity, and political participation: Competing models and contrasting explanations." *The Journal of Politics* 61(4):1092–1114.
- Lewis, Oscar, Victor Barnouw et al. 1958. Village Life in Northern India. Studies in a Delhi Village. PhD thesis Urbana: University of Illinois Press. Illinois, USA.
- Lindberg, Staffan I. 2010. "What accountability pressures do MPs in Africa face and how do they respond? Evidence from Ghana." *The Journal of Modern African Studies* 48(01):117–142.
- Miguel, Edward. 2004. "Tribe or nation? Nation building and public goods in Kenya versus Tanzania." *World Politics* 56(03):328–362.
- Miguel, Edward and Mary Kay Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of public Economics* 89(11):2325–2368.
- Miller, Donald Francis. 1965. "Factions in Indian village politics." *Pacific affairs* 38(1):17–31.
- Min, Brian. 2015. *Power and the Vote: Elections and Electricity in the Developing World*. Cambridge University Press.
- Mishra, Girish and Braj Kumar Pandey. 1996. *Sociology and Economics of Casteism in India: A Study of Bihar*. South Asia Books.
- Montalvo, Jose G and Marta Reynal-Querol. 2005. "Ethnic diversity and economic development." *Journal of Development economics* 76(2):293–323.
- Pande, Rohini. 2003. "Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India." *The American Economic Review* 93(4):1132–1151.

- Shah, Ghanshyam. 2006. *Untouchability in rural India*. Sage.
- Shami, Mahvish. 2012. "Collective action, clientelism, and connectivity." *American Political Science Review* 106(03):588–606.
- Singh, Prerna and Matthias vom Hau. 2016. "Ethnicity in Time Politics, History, and the Relationship between Ethnic Diversity and Public Goods Provision." *Comparative Political Studies* p. 0010414016633231.
- Srinivas, Mysore Narasimhachar. 1987. *The dominant caste and other essays*. Oxford University Press Delhi.
- Thakur, Sankarshan. 2000. *The making of Laloo Yadav: the unmaking of Bihar*. Harper-Collins Publishers India.
- Trounstine, Jessica. 2015. "Segregation and Inequality in Public Goods." *American Journal of Political Science* .
- Tsai, Lily L. 2007. "Solidary groups, informal accountability, and local public goods provision in rural China." *American Political Science Review* 101(02):355–372.
- Whitt, Sam and Rick K Wilson. 2007. "The dictator game, fairness and ethnicity in postwar Bosnia." *American Journal of Political Science* 51(3):655–668.
- Wimmer, Andreas. 2015. "Is Diversity Detrimental? Ethnic Fractionalization, Public Goods Provision, and the Historical Legacies of Stateness." *Comparative Political Studies* p. 0010414015592645.
- Witsoe, Jeffrey. 2013. *Democracy Against Development: Lower-caste Politics and Political Modernity in Postcolonial India*. University of Chicago Press.

Online Appendix

A.1 Measurement

As mentioned in Section Four, the measures of village caste population used here are based on an incomplete (though large) subset of castes, and on household rather than on individual population. Table A.8 tests the robustness of the results to using alternative measures, including using only villages where we have data on the castes of over 80% of individuals, normalizing the caste populations by the total village population rather than the population within the seven enumerated castes, and weighting the number of households for each caste category by the average size of household from that category in the household sample. In nearly all of these models, the effect of upper caste presence on public goods remains significant and positive, while the effect of caste diversity is never negative and occasionally statistically significant and positive. These results provide some indication that the measurement procedure is not driving the results.

Panel D of Table A.8 tests a variant of the diversity hypothesis, which holds that interethnic conflict will be more common in areas where the distribution of ethnic groups is approximately bimodal, or polarized (Montalvo and Reynal-Querol, 2005). However, polarization appears to have little effect on public goods provision, while the effect of upper caste population remains constant.

A.2 Endogenous Identity: Salience, Segregation and Inequality

Salience: Tables A.2 and A.4 test whether the positive effects of upper caste presence are driven by differences in the salience of caste. By and large, the estimated effects of upper caste population in these models remain similar in size to those in Tables 1 and 2 and statistically significant, providing some support for the importance of overall

caste composition rather than the local salience of caste. The first two panels test the effects of differences in ascribed caste status between groups in a village, which might potentially make caste divisions potentially more salient. Panel B is a binary measure of whether the two largest castes in the village are from different status categories (SC, OBC etc.), while Panel A includes a more complex measure of the mean social distance between individuals in a village, with social distance being scored as zero for members of the same caste, one for members of the same caste category, and two for all other pairings. Neither of these measures of differences in group status has a negative effect on public goods provision, and in fact their estimated effect is statistically significant and positive for some categories of goods. In addition, controlling for these factors does not reduce the estimated effect of upper caste residents on provision.

Segregation: Panel C tests the hypothesis that residential segregation effects public goods provision by making ethnic differences more salient and cooperation more difficult. Segregation is operationalized as the proportion of wards (*tolas*) within a village which were either entirely inhabited by one caste or where the two largest castes in the tola were of the same status.⁵ Overall, northern Indian villages are highly segregated by caste status: 68% of tolas in the sample met this criteria. However, with the exception of electricity hours, residentially segregated villages do not appear less likely to have access to public goods than other villages, and the relationship between upper castes and service provision remains strong.

Intergroup Inequality: Tables A.5 and A.6 examine the effect on intergroup ethnic inequality on public goods provision. Panel A is the most direct measure, calculating Generalized Entropy index levels of within and between-caste category inequality in land, using the household survey data. While this measure is imperfect (due to the small number of households surveyed in each village) they provide little evidence for an negative effect. The estimated effect of between-group inequality on public goods is actually positive (though statistically insignificant) and including these measures does not reduce the estimated of upper caste population. Panel B uses a more complex measure of interethnic inequality that uses the whole village population, calculating the

⁵The standard segregation measure, the index of difference, cannot be used for lack of tola-level population data.

rank correlation between castes' proportion of the village population and their proportion of land owned.⁶ even by this measure, interethnic inequality does not appear to have a significant negative effect on public goods provision, and does not reduce the effect of upper castes. In fact, a close association between land and group population size has a substantial estimated negative effect on the provision of schools and health centers.

Total Inequality: Panel C of Tables A.5 and A.6 tests the effect of overall inequality economic within the surveyed households, rather than between group inequality, using the percent of landless households (also used in the main controls) in the village as a rough proxy.⁷ Landlessness has no consistent association with public goods provision, and its inclusion does not change the estimated effect of upper castes. Those interested in the effect of inequality should also note that the overall Generalized Entropy index levels of inequality for the village are a linear combination of the two GE subcomponents in Panel A.

A.3 Local Excludability

While this paper focuses on the consequences of violations of the power and provision assumptions, it is worth noting that the usage of the term “public goods” in the diversity literature is often misleading. In the conventional definition, all local public goods are non-excludable, and once provided within a specific locality are available to all residents of that area. Exclusion thus occurs in the location of goods (i.e. in one village over another) rather than their operation. Local non-excludability is central to many theories of the diversity penalty: Since goods are non-excludable, citizens may wish to free-ride on their provision, a tendency that must be checked through ethnically-based social sanctions. However, many of the services commonly discussed as public goods in the literature are locally excludable in practice, though they may have spillovers beyond the intended recipients (Besley et al., 2004). A village might have an electric or water source, but not allow all households and individual connection (Min, 2015). At the extreme,

⁶The UPBSLS gives the rank of caste groups in order of landownership, but does not give the amounts of land owned or the landowning status of particular individuals. This makes it impossible to calculate land gini coefficients for villages, or decompose within-village inequality.

⁷No data on income is available in the UPBSLS.

individuals can be threatened for making use of otherwise non-excludable goods such as roads, as in some Indian villages (Shah, 2006). The existence of exclusionary practices would seem to be strong evidence that the provision process involves factors other than interethnic cooperation towards a common goal.

An additional implication of a discrimination-based theory is that there will be strong incentives for public goods to be made excludable within local communities. Whether they are acting from prejudice or electoral self-interest, a decision maker intent on favoring local members of the powerful group has little desire for local members of the non-powerful group to take advantage of the good. To the extent the goods are rival (and , additional users from less powerful groups may also degrade the benefits to the targeted group. Even if the targeted group does not lose from having their neighbors share the good, they may prefer to keep the benefit narrow to prevent free riding by the those who do not politically support the decision-maker. Since the diversity hypothesis holds that the benefits of the good should be broadly distributed across groups in order to incentivize cooperative provision, evidence of ethnic exclusion in public goods would seem to cast doubt on the idea that these goods were provided by a process of inter-ethnic cooperation.

Exclusion can thus serve to enhance the efficiency of the discrimination mechanism. One objection to explaining public goods in India as a product of discrimination in favor of the upper castes is that the proportion of upper castes is quite small—only 23% in the 75th percentile village and 3.1% in the median village. Any attempt to distribute to upper caste villages will thus tend to benefit many non-upper caste individuals, and thus be “inefficiently” targeted. However, exclusion provides an explanation for provision to areas where the discriminating group is a minority. Note again that diversity theories have no solid explanation for making public goods excludable, since in these theories the broadly beneficial nature of the good is the primary incentive for cooperation.

Table A.3 presents data on the the degrees to which services tend to benefit the upper castes more than lower castes. A simple but subtle sign of this favoritism is the fact that village facilities tend to be placed in upper caste tolas of a village, making their use both more convenient for upper castes and sometimes allowing up caste members to intimidate those from other parts of the village who attempt to take advantage of

the service (Shah, 2006). Within the villages with primary schools in our sample, the average upper caste household had a 9.6 minute walk to the nearest a primary school, as opposed to 12.7 minutes for members of other groups. This pattern of uneven spatial distribution was also reflected in the laying of power lines. Conditional on the village having electricity, 100% of upper caste neighborhoods in the sample (coded by the largest group) had electricity, while only 40% of other neighborhoods did so. Another way of turning public goods into private benefits is to make sure their staffs are made up of members of one's own group (Chandra, 2007). Despite a longstanding policy of ethnic quotas in government hiring, the UPBSLS sample, the Anganwadi (child health care) worker and the shopkeepers of government ration shops, are also disproportionately upper caste.

Table A.1: Summary statistics

Variable	Mean	Std. Dev.	N
Prop. Segregated Tolas	0.687	0.234	124
Village Pop.	257.283	163.68	120
Total HH Exp.	6531.271	3233.422	120
HH Asset Factor Score	0.01	0.356	120
Quality Housing	0.186	0.178	120
Village Prop. Upper Castes	0.137	0.2	120
Land-Population Rank Correlation	0.819	0.08	120
Avg. Upper Caste Land-Avg. Other Caste Land	3.035	5.55	75
Village Prop. Landless	23.575	20.932	120
Village Top Two Castes of Different Status	0.806	0.397	124
Total Village Social Distance	0.568	0.207	124
Village Caste Fractionalization	0.646	0.197	120
Village Prop. Scheduled Castes/Tribes	0.274	0.208	120
Village Prop. Muslim	0.119	0.227	120
% House. With Off-Farm Jobs	39.586	25.881	116
Village Road Access	0.895	0.308	124
Irrigation % Scoring	3.758	1.188	120
Avg. Price Non-Irrigated Land	54688.333	37894.802	120
Village Prop. Landless* Fractionalization	15.465	14.982	120
GE Between Group Inequality	0.238	0.279	120
GE Within Group Inequality	0.286	0.162	120

Table A.2: Diversity and Social Difference

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Social Distance</i>				
Total Village Social Distance	0.197 (2.583)	3.782* (1.934)	-0.929 (1.356)	8.985** (4.073)
Village Pop.	0.00350 (0.00287)	0.00172 (0.00130)	0.0108*** (0.00307)	0.00339** (0.00159)
Constant	3.149* (1.627)	-3.981*** (1.268)	-0.279 (0.881)	-8.897*** (2.834)
Observations	120	119	119	119
<i>Panel B: Status Differences Among Large Groups</i>				
Village Top Two Castes of Diff. Status	-0.752 (1.232)	0.236 (0.615)	1.176* (0.610)	0.221 (0.835)
Village Pop.	0.00360 (0.00282)	0.00219* (0.00125)	0.00986*** (0.00293)	0.00404*** (0.00151)
Constant	3.867*** (1.319)	-1.970*** (0.671)	-1.604** (0.736)	-3.410*** (0.939)
Observations	120	119	119	119
<i>Panel C: Within Village Segregation</i>				
Prop. Segregated Tolas	-4.611** (2.124)	-0.351 (0.982)	1.308 (1.318)	-1.654 (1.209)
Village Pop.	0.00107 (0.00299)	0.00202 (0.00135)	0.0118*** (0.00345)	0.00332** (0.00164)
Constant	7.010*** (1.924)	-1.491* (0.886)	-1.925 (1.331)	-1.979* (1.057)
Observations	120	119	119	119

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

Table A.3: Discrimination in Service Provision

MEASURE	Upper Castes	Other Groups	N
<i>Panel A: Relative Service Levels: “Public” Goods</i>			
Avg. Walking Time to Primary School	9.6	12.7	95
Prop. Home Electricity	.29	.16	55
Prop. Children Immunized	.87	.83	25
Prop. Children Attending School	.75	.51	95
Prop. Caste Neighborhood with Electricity	1	.40	55
Prop. Caste Neighborhood with High Primary Enrollment	.69	.55	95
<i>Panel B: Caste of Administrators</i>			
Population Overall	.136	.864	120
Pradhan	.206	.794	120
Anganwadi Worker	.414	.586	41
PDS Shopkeeper	.295	.705	71

Table A.4: Social Difference and Social Power

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Social Distance</i>				
Total Village Social Distance	-0.178 (2.562)	3.822* (1.952)	-1.458 (1.488)	8.810** (4.047)
Village Pop.	0.00391 (0.00284)	0.00190 (0.00133)	0.0126*** (0.00337)	0.00342** (0.00161)
Village Prop. Upper Castes	4.357* (2.291)	2.497** (1.096)	3.921** (1.897)	2.496 (1.655)
Constant	2.670 (1.629)	-4.461*** (1.320)	-0.747 (0.960)	-9.216*** (2.856)
Observations	120	119	119	119
<i>Panel B: Status Differences Among Large Groups</i>				
Village Top Two Castes of Diff. Status	-1.243 (1.238)	-0.0183 (0.639)	0.849 (0.647)	0.0112 (0.863)
Village Pop.	0.00400 (0.00278)	0.00243* (0.00127)	0.0113*** (0.00327)	0.00419*** (0.00152)
Village Prop. Upper Castes	4.792** (2.318)	2.313** (1.036)	2.945* (1.687)	2.143 (1.314)
Constant	3.519*** (1.312)	-2.186*** (0.696)	-1.955** (0.792)	-3.636*** (0.978)
Observations	120	119	119	119
<i>Panel C: Within Village Segregation</i>				
Prop. Segregated Tolas	-4.542** (2.101)	-0.344 (1.008)	1.479 (1.357)	-1.672 (1.228)
Village Pop.	0.00143 (0.00296)	0.00225 (0.00137)	0.0136*** (0.00376)	0.00347** (0.00165)
Village Prop. Upper Castes	4.261* (2.240)	2.308** (1.016)	3.528** (1.692)	2.170* (1.314)
Constant	6.288*** (1.940)	-1.926** (0.932)	-2.786* (1.428)	-2.377** (1.111)
Observations	120	119	119	119

Standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

Table A.5: Diversity and Inequality

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Between Group Land Inequality</i>				
GE Between Group Inequality	1.525 (1.655)	0.937 (0.725)	0.247 (1.020)	0.361 (1.007)
GE Within Group Inequality	3.277 (2.845)	-0.849 (1.419)	-1.369 (1.545)	-2.265 (2.077)
Village Pop.	0.00355 (0.00282)	0.00238* (0.00126)	0.0110*** (0.00316)	0.00419*** (0.00153)
Constant	1.952 (1.263)	-1.824*** (0.611)	-0.512 (0.718)	-2.748*** (0.839)
Observations	120	119	119	119
<i>Panel B: Land Inequality (Rank Correlation)</i>				
Land-Population Rank Correlation	0.766 (5.782)	0.629 (2.785)	-6.951* (3.548)	-7.285** (3.618)
Village Pop.	0.00353 (0.00282)	0.00219* (0.00125)	0.0123*** (0.00342)	0.00414*** (0.00153)
Constant	2.631 (4.791)	-2.288 (2.328)	4.647* (2.805)	2.617 (2.869)
Observations	120	119	119	119
<i>Panel C: Landlessness</i>				
Village Caste Fract.	-2.331 (3.460)	2.418 (2.705)	-1.285 (1.810)	11.18 (8.713)
Village Pop.	0.00431 (0.00293)	0.00140 (0.00133)	0.0107*** (0.00310)	0.00291* (0.00160)
Village Prop. Landless	-0.127* (0.0682)	-0.0185 (0.0622)	-0.00501 (0.0421)	0.00363 (0.165)
Village Prop. Landless* Fract.	0.155 (0.102)	0.0461 (0.0882)	0.0204 (0.0658)	0.0209 (0.220)
Constant	5.158** (2.229)	-3.525* (1.878)	-0.167 (1.169)	-11.52* (6.594)
Observations	120	119	119	119

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

Table A.6: Inequality and Social Power

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Between Group Inequality</i>				
GE Between Group Inequality	1.001 (1.673)	0.704 (0.736)	-0.333 (1.075)	0.145 (1.028)
GE Within Group Inequality	2.661 (2.849)	-1.355 (1.464)	-1.976 (1.649)	-2.734 (2.132)
Village Pop.	0.00382 (0.00280)	0.00261** (0.00129)	0.0129*** (0.00349)	0.00432*** (0.00154)
Village Prop. Upper Castes	3.854 (2.340)	2.382** (1.057)	3.590** (1.712)	2.500* (1.387)
Constant	1.658 (1.267)	-2.067*** (0.632)	-0.970 (0.783)	-3.022*** (0.878)
Observations	120	119	119	119
<i>Panel B: Land Inequality (Rank Correlation)</i>				
Village Pop.	0.00387 (0.00280)	0.00242* (0.00127)	0.0143*** (0.00379)	0.00440*** (0.00153)
Village Prop. Upper Castes	4.345* (2.290)	2.302** (1.017)	3.653** (1.695)	2.705** (1.358)
Land-Population Rank Correlation	0.00380 (5.733)	0.197 (2.924)	-7.522** (3.621)	-8.765** (3.909)
Constant	2.574 (4.738)	-2.361 (2.432)	4.322 (2.830)	3.255 (3.016)
Observations	120	119	119	119
<i>Panel C: Landlessness</i>				
Village Pop.	0.00445 (0.00283)	0.00216* (0.00129)	0.0118*** (0.00328)	0.00389** (0.00154)
Village Prop. Landless	-0.0242 (0.0222)	0.0162 (0.0109)	0.00852 (0.0137)	0.0208 (0.0142)
Village Prop. Upper Castes	4.084* (2.285)	2.546** (1.041)	3.464** (1.655)	2.536* (1.355)
Constant	3.035*** (1.008)	-2.570*** (0.559)	-1.590** (0.706)	-4.156*** (0.819)
Observations	120	119	119	119

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

Table A.7: District FE and Simple Models

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: District FE</i>				
Village Pop.	0.0110*** (0.00296)	0.00454** (0.00227)	0.0161*** (0.00503)	0.00730** (0.00303)
Village Prop. Upper Castes	1.077 (2.236)	4.042** (1.801)	5.304** (2.598)	4.065 (2.725)
Constant	5.070** (1.952)	-1.620 (1.073)	-1.884 (1.491)	-4.051*** (1.487)
District FE	Yes	Yes	Yes	Yes
Observations	120	81	79	54
<i>Panel B: Simple Model</i>				
Village Caste Fractionalization	2.093 (2.317)	4.964** (2.104)	0.353 (1.116)	14.03*** (5.166)
Village Prop. Upper Castes	4.180* (2.281)	2.893** (1.151)	2.362 (1.536)	3.683** (1.782)
Constant	2.243 (1.599)	-5.051*** (1.594)	0.833 (0.739)	-12.86*** (4.055)
Observations	120	119	119	119

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

Table A.8: Measurement of Caste Populations

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Only High Information Villages</i>				
Village Pop.	0.00525 (0.00319)	0.00275* (0.00153)	0.0112*** (0.00351)	0.00358* (0.00192)
Village Prop. Upper Castes	3.375 (2.408)	2.567** (1.250)	3.946** (1.957)	5.051** (1.964)
Village Caste Fractionalization	1.804 (2.465)	4.327* (2.250)	-1.062 (1.342)	9.977** (4.923)
Constant	1.238 (1.703)	-5.283*** (1.662)	-0.738 (0.954)	-11.27*** (3.880)
Observations	100	99	99	99
<i>Panel B: Proportions of Total Population</i>				
Village Pop.	0.00349 (0.00286)	0.00157 (0.00140)	0.0128*** (0.00345)	0.00267 (0.00168)
Village Prop. Upper Castes	3.737 (2.253)	2.591** (1.151)	3.397** (1.725)	3.037* (1.625)
ELF	1.170 (2.167)	3.688* (1.904)	-0.998 (1.224)	7.336** (3.624)
Constant	1.836 (1.577)	-4.769*** (1.467)	-0.960 (0.918)	-8.901*** (2.916)
Observations	117	116	116	116
<i>Panel C: Proportions Weighted by Household Size</i>				
Village Pop.	0.00349	0.00172	0.0122***	0.00316*
Weighted Prop. Upper Castes	3.734* (2.177)	2.617** (1.103)	3.142** (1.544)	4.250** (1.788)
Weighted ELF	1.310 (2.365)	4.113** (2.059)	-0.843 (1.284)	13.12** (5.132)
Constant	1.854 (1.651)	-4.919*** (1.530)	-0.964 (0.929)	-13.34*** (4.067)
Observations	120	119	119	119
<i>Panel D: Polarization</i>				
Village Pop.	0.00393 (0.00281)	0.00217* (0.00132)	0.0124*** (0.00334)	0.00399** (0.00156)
Village Prop. Upper Castes	4.345* (2.284)	2.694** (1.101)	3.715** (1.817)	2.432* (1.435)
Polarization	-0.194 (0.965)	1.328* (0.737)	-0.442 (0.547)	1.057 (0.979)
Constant	1.742 (4.253)	3.306 (3.039)	-3.397 (2.518)	0.761 (4.048)
Observations	120	119	119	119
R-squared	0.044			

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable. Panel A examines only villages for whom we have information on the caste of more the 90% of inhabitants. Panel B calculates the proportions of caste based on the total population rather than the population of the seven largest castes. Panel C calculates the proportions of caste using the proportion of households weighted by the average household size within each caste category.

Table A.9: Connections and Social Capital

MEASURE	Upper Castes	Other Groups	N
<i>Panel A: Government Connections</i>			
Social Network Gov. Employee	.6	.27	2727
Social Network Gov. Officer	.24	.08	2727
Social Network Gov. Emp. Same Caste	.62	.37	1095
Social Network Gov. Emp. Outside Village	.33	.22	2727
<i>Panel B: Associational Membership</i>			
Caste Association Member	.09	.16	2763
Other Association Member	.19	.21	2763
Gov. Cooperative Member	.05	.02	2763
Agree "All should work together to Solve Local Problems"	.48	.5	2763

Taken from the Indian Human Development Survey, 2005-6, using only Northern and Central Bihar and Eastern UP districts.

Table A.10: Wealth and Public Goods in India

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Monthly Expenditures</i>				
Village Pop.	0.00356 (0.00287)	0.00168 (0.00135)	0.0124*** (0.00337)	0.00297* (0.00162)
Village Prop. Upper Castes	3.997* (2.342)	2.784** (1.192)	3.488* (1.813)	3.691* (1.956)
Village Caste Fractionalization	1.303 (2.386)	4.258** (2.116)	-0.997 (1.302)	13.21** (5.498)
Total HH Exp.	9.99e-05 (0.000145)	3.16e-05 (7.23e-05)	3.62e-05 (7.63e-05)	-1.54e-05 (0.000106)
Constant	1.211 (1.821)	-5.209*** (1.571)	-1.121 (1.041)	-13.06*** (4.243)
Observations	120	119	119	119
<i>Panel B: Household Assets</i>				
Village Pop.	0.00488* (0.00271)	0.00177 (0.00136)	0.0126*** (0.00353)	0.00348** (0.00169)
Village Prop. Upper Castes	0.718 (2.323)	2.519** (1.239)	5.397*** (2.066)	2.734 (1.893)
Village Caste Fractionalization	1.025 (2.230)	4.509** (2.184)	-1.140 (1.448)	14.77*** (5.541)
HH Asset Factor Score	5.333*** (1.318)	0.606 (0.672)	-1.914** (0.815)	1.954* (1.005)
Constant	2.101 (1.552)	-5.178*** (1.612)	-0.944 (1.032)	-14.48*** (4.409)
Observations	120	119	119	119
<i>Panel C: Housing</i>				
Village Pop.	0.00570** (0.00260)	0.00169 (0.00136)	0.0125*** (0.00342)	0.00374** (0.00175)
Village Prop. Upper Castes	1.663 (2.111)	2.857** (1.177)	3.987** (1.793)	3.685** (1.807)
Village Caste Fractionalization	1.555 (2.132)	4.413** (2.115)	-1.155 (1.360)	12.96** (5.222)
Quality Housing	12.82*** (2.402)	0.334 (1.404)	-1.323 (1.409)	3.363* (1.994)
Constant	-0.917 (1.564)	-5.183*** (1.604)	-0.595 (1.018)	-13.94*** (4.214)
Observations	120	119	119	119

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.

Table A.11: Clientelism as an Alternative Mechanism

VARIABLES	(1) Elect. Hours	(2) Telephone	(3) Pub. Prim. School	(4) Health Center
<i>Panel A: Lending</i>				
Village Pop.	0.00432 (0.00279)	0.00226* (0.00128)	0.0120*** (0.00329)	0.00446*** (0.00160)
Village Prop. Upper Castes	2.552 (2.512)	2.815** (1.118)	3.254* (1.840)	2.702** (1.340)
Prop. Upper Caste Lender to Lower Castes	1.193 (1.902)	0.0958 (0.917)	-0.0125 (1.061)	-2.222 (1.508)
Constant	2.312** (1.018)	-2.238*** (0.533)	-1.425** (0.715)	-3.250*** (0.703)
Observations	119	118	118	118
R-squared	0.034			
<i>Panel B: Casual Workers</i>				
Village Pop.	0.00380 (0.00275)	0.00246* (0.00128)	0.0122*** (0.00332)	0.00419*** (0.00152)
Village Prop. Upper Castes	5.721** (2.373)	2.677** (1.100)	3.264* (1.712)	2.154 (1.372)
Lower Caste Prop. Causal Labor	-4.718* (2.566)	-1.202 (1.251)	0.807 (1.455)	-0.0262 (1.648)
Constant	5.152*** (1.669)	-1.570** (0.796)	-1.909* (1.084)	-3.613*** (1.110)
Observations	120	119	119	119
R-squared	0.070			

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Column 1 shows the results of linear regression models with the number of hours of electricity a village gets in a “good” month as the dependent variable. Columns 2-4 are logistic regressions with a binary measure of whether a village has a particular facility as the dependent variable.